ol. 31, No. 2 April, 1947

Journal of Applied Psychology

EDITED BY: DONALD G. PATERSON, UNIVERSITY OF MINNESOTA

Consulting Editors

PAUL S. ACHILLES, *Psychological Corporation;* WALTER V. BINGHAM, *A.G.O., War Department;* HAROLD E. BURTT, *Ohio State University;* ARTHUR I. GATES, *T. C. Columbia University;* JOHN G. JENKINS, *University of Maryland;* IRVING LORGE, *T. C. Columbia University;* QUINN MCNEMAR, *Stanford University;* WILLARD C. OLSON, *University of Michigan;* JAMES P. PORTER, *Swarthmore, Pennsylvania;* EDWARD K. STRONG, JR., *Stanford University;* MORRIS S. VITELES, *University of Pennsylvania;* JOSEPH ZUBIN, *N. Y. Psychiatric Institute.*

Table of Contents

Job Attitudes of Life Insurance Agents: S. HABBE 111
The Psychological Barometer of Public Attitudes: H. C. LINK 129
How Many Foot Candles?: C.-E. A. WINSLOW 140
Some Comments on the Army General Classification Test: A. J. DUNCAN 143
The MacQuarrie Test for Mechanical Ability: II. Factor Analysis: C. H. GOODMAN 150
Reliability of the Grove Modification of the Kent-Shakow Formboard Series: R. C. WYLIE 155
Influence of Leading Upon Readability of Newspaper Type: D. G. PATERSON AND M. A. TINKER 160
Prediction of Success in Learning Japanese: S. B. WILLIAMS AND H. J. LEAVITT 164
Effect of a Guidance Program on Emotional Development: G. M. WORBOIS ... 169
Further Application of Simplified Scoring of the Bernreuter Personality Inventory: D. C. MCCLELLAND 182
Braille Edition of the Minnesota Multiphasic Personality Inventory for Use with the Blind: O. H. CROSS 189
Studies in Projective Techniques: I. The Reliability of a Multiple Choice Group Rorschach Test: C. H. LAWSHE, JR., AND M. H. FORSTER 199
A Note on the Hankes Test Scoring Machine: E. K. STRONG, JR., AND E. J. HANKES 212
Book Reviews 215
New Books, Monographs, and Pamphlets 222

Published Bi-monthly by The American Psychological Association, Inc.
Prince and Lemon Sts., Lancaster, Pa., and
1515 Massachusetts Ave., NW, Washington 5, D. C.

Entered as second-class matter, August 19, 1943, at the post office at Lancaster, Pa., under the Act of March 3, 1879

Copyright, 1947, by The American Psychological Association, Inc.

PUBLICATIONS OF

The American Psychological Association, Inc.

PSYCHOLOGICAL REVIEW

Editor: HERBERT S. LANGFELD, *Princeton University*

Contains original contributions of a theoretical nature; bi-monthly.
Subscription: $5.50 (Foreign $5.75). Single copies, $1.00.

PSYCHOLOGICAL BULLETIN

Editor: LYLE H. LANIER, *Vassar College*

Contains critical reviews of books and articles and critical and analytical summaries of psychological fields or subject matter; bi-monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

PSYCHOLOGICAL MONOGRAPHS

Editor: JOHN F. DASHIELL, *University of North Carolina*

Contains longer researches and laboratory studies which appear as units; published at irregular intervals at a cost to author of about $2.50 a page; author receives 150 copies gratis.
Subscription: $6.00 per volume of about 350 pages (Foreign $6.30). Single copies, price varies according to size.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

Editor: GORDON W. ALLPORT, *Harvard University*

Contains original contributions in the field of abnormal and social psychology, reviews, and case reports; quarterly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Editor: FRANCIS W. IRWIN, *University of Pennsylvania*

Contains original contributions of an experimental character; bi-monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

Editor: C. M. LOUTTIT, *Sampson College*

Contains noncritical abstracts of the world's literature in psychology and related subjects; monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $.75.

JOURNAL OF APPLIED PSYCHOLOGY

Editor: DONALD G. PATERSON, *University of Minnesota*

Contains material covering applications of psychology to business, industry, education, etc.; bi-monthly.
Subscription: $6.00 (Foreign $6.30). Single copies, $1.25.

APPLIED PSYCHOLOGY MONOGRAPHS

Editor: HERBERT S. CONRAD, *College Entrance Examination Board, Princeton*

Contains longer researches and studies in the field of applied psychology; published at irregular intervals at a cost to author of about $2.00 a page.
Single copies only, price varies according to size.

JOURNAL OF CONSULTING PSYCHOLOGY

Editor: LAURANCE F. SHAFFER, *Teachers College, Columbia University*

Contains articles in the field of clinical and consulting psychology, counseling and guidance; bi-monthly.
Subscription: $3.00 (Foreign $3.50). Single copies, $.60.

THE AMERICAN PSYCHOLOGIST

Editor: DAEL WOLFLE, *American Psychological Association*

Contains all official papers of the Association and articles concerning psychology as a profession; monthly.
Subscription: $7.00 (Foreign $7.50). Single copies, $.75.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY

Editor: CALVIN P. STONE, *Stanford University*

Contains original contributions in the field of comparative and physiological psychology; bi-monthly.
Subscription: $7.00 (Foreign $7.50). Single copies, $1.25.

Subscriptions are payable in advance and are terminated at expiration.
Make checks payable to the American Psychological Association, Inc.

Subscriptions, orders, and other business communications should be sent to:

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 MASSACHUSETTS AVENUE N.W., WASHINGTON 5, D.C.

Science
Psychology
Dept
2-10-49
Shortage

Journal of Applied Psychology

Vol. 31, No. 2 April, 1947

Job Attitudes of Life Insurance Agents

Stephen Habbe

Life Insurance Agency Management Association, Hartford, Connecticut

Contrary to what many persons would suspect, life insurance agents express considerable satisfaction with their jobs.

Surveys of job preference among college students and other groups have shown life insurance selling to be one of the least popular of all occupations. Rarely does a college student, nearing graduation, announce his intentions of seeking a career as an insurance agent. The situation is not far different among older adults who have been in the labor market some years and who, presumably, are in a better position than college students to appraise different job opportunities. Thus, the figures which will be presented in the following pages will come as a surprise to some readers.

The Life Insurance Agency Management Association, successor to the Life Insurance Sales Research Bureau, began the study of the job satisfaction of life insurance agents late in 1943. Work proceeded for over a year during which period the research forms and techniques were being revised and improved. In March, 1945, a progress report, "The Job Satisfaction of Life Underwriters," was published. This report described the project and analyzed the findings from the experimental surveys of agents in five insurance companies. As a result of this publication and of talks and discussions among insurance executives around the country, interest was aroused in the job satisfaction project. No fewer than seventeen surveys have now been made and others are scheduled. Several companies already have asked for repeat surveys among their agents, and it is planned to make these every three to five years, wherever desired. The Agency Management Association recently developed a Job Opinion form for life insurance General Agents (Agency Managers), and it is expected that research with this group will proceed concurrently with the Job Satisfaction studies among agents. Finally, a few companies have become interested in surveying the job attitudes of their Home Office employees.[1]

[1] There are over 350 life insurance companies domiciled in the United States. All have central and field organizations. Some carry on very small operations and may

How A Survey Is Made

Let us suppose that the fictitious Fireside Life Insurance Company wishes to survey the job attitudes of its field agents. It communicates its desire to the Agency Management Assoc. of which it is a member, explaining that it has a total of 300 agents under contract in 40 agencies in 28 states. The Association immediately supplies the company with the necessary materials for a survey. This means a printed questionnaire form and a business reply envelope for each agent. Also, the Association explains the mechanics of a survey.

As a result of its earlier experiences, the Association is able to "guarantee" Fireside Life an eighty to ninety per cent return of questionnaires, providing its suggestions and instructions are followed. A pattern of effective procedures has been worked out, and in this pattern there are several important points:

1. Three letters are written each Agency Manager by the Fireside's Agency Vice President. The first letter explains the Job Satisfaction project and advances reasons why the company is having a survey made. Cooperation of each field office is assumed—and usually secured—but no coercion is exercised if a Manager decides not to distribute the questionnaires among his agents. A second letter follows a week later. This letter explains how the survey is conducted. The questionnaires and prepared envelopes are enclosed with the second letter. A brief third letter from the Vice President arrives about four or five days after the second letter. It urges immediate followup of "delinquent" agents and all possible effort to obtain 100% cooperation within the agency office.
2. The "confidential mediator" technique is used. Each Sales Manager is asked to make a list of his agents. Also, he is asked to distribute the Job Satisfaction materials personally to each agent. Finally, it is suggested that he instruct each agent to notify him when his blank has been completed and mailed. Thus, the Manager has a constant check on the cooperation of his agents and after a few days he can follow up the names not yet crossed off the list.
3. An eighty-word "sales talk" is provided each Manager. This is a part of the second letter from the Home Office. It is a simple, straightforward explanation of the survey and a request for the agent's cooperation.

Evidence of Agents' Interest

Most life insurance agents are in an anomalous position—they are not company employees in the usual sense, nor are they strictly independent business men. They like freedom from close, petty supervision and yet they have a deep feeling of loyalty to their company. Most of them are at a distance—psychological as well as geographical—from their Home Offices and it is unlikely that they see the "Big Bosses" more often than once a year.

do business in only one state. At the other extreme is a company employing some 40,000 persons, which has been described as the world's largest private business. The 155 U. S. companies, which are members of the Agency Management Association, write over 90% of the insurance in this country.

The majority of the 9,353 agents, presumably receiving blanks[2] in the 17 companies already surveyed, seemed interested in this research. Perhaps the best evidence of this is the fact that 6,992 of the questionnaires were filled out and returned to the Agency Management Association, an over-all return of 75%. Individual company returns varied from a low of 64% to a high of 96%. The average return for the five companies surveyed early in 1946, is 86%, reflecting the improved procedures adopted during the later stages of the work.

Another evidence of interest is the care and thoroughness with which the blanks are completed. Over 90% of the blanks returned by agents are filled out correctly and are usable. Many agents go to special efforts, printing or typing their answers. More than a few, entirely on their own initiative, add extra pages of comments and send these along with their blanks.

Finally, a number of the agents write comments of direct appreciation on the margins of their blanks. A few of these follow: "This is a hell of a good idea"!; "More power to you"; "I've sold for my company for 18 years and this is the first time anyone has ever consulted with me about what I thought"; "You get the feeling out here that you're a long ways from home base, and doing something like this makes you feel your ideas count and are worth something to others." Although the questionnaire form is anonymous and no place is provided for signatures, quite a few agents write their names and addresses on their blanks and some go further and enter into correspondence with the Management Association.

Purposes

Let us suppose that the survey for Fireside Life has been completed. The opening paragraph in the report would read something like this: "This study of Job Satisfaction in the Fireside Life Insurance Company was made during September, 1946, in order to see what could be learned about the attitudes of full-time agents toward the institution of life insurance, toward the company they work for, and toward various aspects of their job. It is hoped that the results of this and similar studies will not only assist the company to improve the situation of its own men, but will also develop information significant to the business as a whole."

This is a statement of over-all purpose. A company may be interested in learning a number of specific things about its operations. What, for example, do its fulltime fieldmen think of the company's retirement

[2] It cannot be said that 9,353 agents *certainly* received questionnaire forms. Sick agents, agents on vacation, rural agents at a distance from the field office, etc. in some cases undoubtedly failed to receive blanks. In a few instances Managers failed to distribute *any* of the blanks sent to them for their agents.

plan?[3] When the returns are in, the company has the answer. The agents feel that the retirement plan needs to be strengthened. A number comment in detail that the retirement plan should be built primarily with the career underwriter (agent) in mind. The company considers these findings and suggestions. It gathers information from other sources, too, because other things are frequently as important as what agents want. Finally, with all the facts at hand the company decides what should be done.

Other important purposes which a Job Satisfaction survey may serve include: 1. To give recognition to each full-time underwriter by asking his opinions about a number of field problems. This, in itself, is a morale-building step if the company follows through and acts on the findings; 2. To furnish management aids. Management, like the individual, often has blind spots. Management *thinks* it knows what the workers want. Often it is right, but sometimes its judgments miss the mark. When decisions are to be made that affect workers, it is well to get their attitudes toward them beforehand; and 3. To indicate and appraise the general field situation by a comparison of company scores with the average scores of other companies.

How Many Workers Are Satisfied?

Hoppock's summaries of the results of job satisfaction studies made by different workers in different occupations indicate that about three workers in four are reasonably content in their jobs.[4] For scientific purposes this statement would require considerable explanation, but for the present purpose we are concerned only with establishing a general level of workers' satisfaction. Let us say that 75% is a fair estimate, recognizing that that satisfaction indices vary greatly by occupation and that they vary greatly, too, according to the different survey instruments used.

There are no standard, accepted forms for testing job satisfaction. Each worker tends to construct his own form, asking those questions which seem most pertinent and most important in the job or in the company to be surveyed. This was the procedure adopted in the present research.

Approximately 100 questions were developed originally. These were field tested. About half the items were retained and several new ones added. Next, a four-page questionnaire form was printed and tried out on an experimental group of five companies, a total of 1,158 agents co-operating. Again, the findings were subjected to critical analysis; a few minor revisions made; and a "final" questionnaire printed. For the

[3] One of the 44 items which appears in the questionnaire.

[4] Hoppock, R. *Job satisfaction.* New York: Harper and Brothers, 1935. Pp. 303.

most part, the data from the first five companies are comparable with the data developed later and thus it was possible to include the 1,158 records in many of the tables which follow.

It is impossible to say whether the present questionnaire reflects a faithful picture of the job attitude of the life insurance agent. Like the mirrors at Coney Island, designed to flatter thin persons, the questionnaire may give back too rosy a reflection. On the other hand, perhaps it does the opposite: it is conceivable that the agent in reality is *more* satisfied than the questionnaire answers indicate. In constructing the questionnaire the attempt was made to include questions on all important phases of the agent's job without regard to any estimate of the *favorableness* of the questions. It was felt that the answers to the different items would be more valuable than an over-all, composite satisfaction score.

This has been the case in actual experience, and no great significance is attached to the fact that 79% of the agents surveyed to date rate themselves as satisfied with their job. The figure appears good, and there is rationale to support it, but other figures are more important.

Description of the Questionnaire

The first page of the four-page printed questionnaire is a persuasive letter of explanation addressed to the agent and signed by the Managing Director of the Association. It asks the agent for straight-from-the-shoulder answers and promises that his confidences will be safeguarded. See Figure 1.

The questionnaire proper begins on page 2. The directions are simple as may be seen in Figure 2.

The 44 questions form the main body of the questionnaire, extending to the middle of the third page. The remaining half of this page asks six special questions which are answered differently from the 44 questions. See Figure 3.

Two questions are on page 4 as shown in Figure 4. At the bottom of page 4 are spaces to indicate age ("under 40" or "40 or over"), experience as an agent ("under 2 years" or "2 years or over"), and last year's production ("under $200,000" of insurance or "$200,000 or over"). The average agent is able to complete his blank within 30 minutes.

Main Questionnaire Findings

The 44 questions in the main body of the questionnaire ask, "How do you feel about. . . ?" Forty-three of the questions asked about single phases of the agent's job. One question, inserted inconspicuously between two public relations questions, asked, "How do you feel about your job as a life underwriter?" The purpose of asking this general job

confidential

LIFE INSURANCE AGENCY MANAGEMENT ASSOCIATION

JOHN MARSHALL HOLCOMBE, JR.
MANAGING DIRECTOR
ELIZABETH C. STEVENS
SECRETARY-TREASURER

115 BROAD STREET
HARTFORD 5, CONNECTICUT
WENDELL P. HANSELMAN, PRESIDENT

LEWIS W. S. CHAPMAN, C. L. U.
DIRECTOR OF SERVICE AND COMPANY RELATIONS
CHARLES J. ZIMMERMAN, C. L. U.
DIRECTOR OF INSTITUTIONAL RELATIONS

Dear Underwriter:

You are invited to join with your Company and with the Life Insurance Agency Management Association in an important study of Job Satisfaction. We hope and believe that the results of this survey will benefit the life insurance fraternity, particularly the individual underwriter by adding to the satisfaction he finds in his job.

There are only two things you need to do. First, fill out the remaining pages of this form. Second, send your completed folder direct to the Management Association in the envelope provided. We will do the rest!

The survey is anonymous from start to finish. Only your Company and your Agency are identified. No one apart from the Management Association staff will have access to your folder. The Management Association has a reputation as a professional agency and you can rely upon it to safeguard your confidence.

Your Company has gone over this study carefully and approved it. Some organizations might shy away from such a study, but your Company wants to know what its underwriters are thinking so that it can help them in all possible ways. The Management Association will furnish your Company with a composite report of its findings.

As you will note from our letterhead, the Management Association represents a merger of two agency organizations. It now serves 150 life insurance companies, including your Company. Its purpose is to promote the better distribution of life insurance, and it is with this purpose in mind that this survey is being made. We invite your help and suggestions.

Yours very sincerely,

JMH:ETG
Enc.

Managing Director

FIG. 1. Front page of Job Satisfaction Questionnaire.

JOB SATISFACTION

Explanation

Here are 44 questions which ask you how you feel about your job.

Please answer each question with a number from 5 to 1. The key below will show you how to answer.

Please answer all the questions even though some do not seem to "fit" your case exactly.

Key for Answering the Questions

5 means you are **Extremely Satisfied**
4 means you are **Satisfied**
3 means you have **Balanced Feelings**
2 means you are **Dissatisfied**
1 means you are **Extremely Dissatisfied**

		Answers
How do you feel about . . .		
General Questions	1. The way Life Insurance performs its basic functions?	.....
	2. The public relations programs of Life Insurance companies?	.....
	3. The cost of Life Insurance to the policyholder?	.....
	4. Your job as a life underwriter?	.....
	5. Underwriters of other Life Insurance companies?	.....
	6. The attitude of the public toward Life Insurance underwriters?	.....
	7. The attitude of your family toward your selling life insurance?	.....
How do you feel about your Company's . . .		
Your Company	8. General reputation and public standing?	.....
	9. Underwriting practices?	.....
	10. Desire and willingness to make improvements in Agency matters?	.....
	11. Competitive position in the Life Insurance business?	.....
	12. Interest in the underwriter as an individual?	.....
	13. Cooperation with your Agency in helping underwriters?	.....
	14. Sales promotional material?	.....
How do you feel about . . .		
Your Agency	15. Your Agency Manager?	.....
	16. The other underwriters in your office?	.....
	17. The cashier in your office?	.....
	18. The clerical employees in your office?	.....
	19. Your Agency's location and work surroundings?	.....
	20. Your Agency's willingness to see you through adversities?	.....
How do you feel about . . .		
Conditions and Nature of Work	21. The hours of work your job requires?	.....
	22. The effects of your work on your health?	.....
	23. Preparing briefs and doing other office jobs?	.....
	24. Prospecting?	.....
	25. Interviewing prospects?	.....
	26. Servicing your own policyholders?	.....
	27. Servicing orphan policyholders?	.....
	28. The opportunities in your work for performing useful service?	.....
	29. The opportunities in your work for interesting experiences?	.....

FIG. 2. Second page of Job Satisfaction Questionnaire.

How do you feel about the . . . .

		Answer
Training and Supervision	30. Training you received when you entered the business?	____
	31. Leadership and inspiration your Manager gives you?	____
	32. Help you get in improving your sales techniques in the office?	____
	33. Help you get in improving your sales techniques in the field?	____
	34. Recognition your Manager gives your ability and effort?	____
	35. Fairness and impartiality he shows in his dealings with his men?	____

How do you feel about . . .

Compensation	36. Your earnings so far this year?	____
	37. The regularity of your income?	____
	38. The commission form of compensation?	____
	39. Your Company's present compensation plan?	____

How do you feel about . . .

Future	40. Your probable future earnings as an underwriter?	____
	41. The probable future market for Life Insurance?	____
	42. The opportunities available to you for professional growth? (C. L. U. courses, Company schools, Underwriters' Associations, etc.)	____
	43. The security of your job?	____
	44. Your Company's retirement plan?	____

A. Do you feel the need for more training? ☐₁ Yes ☐₂ No

B. Do you feel the need for closer direction of your activities? ☐₁ Yes ☐₂ No

C. Put a check in one of the boxes below to show how you feel about your present job:

☐₁ Satisfied with it 0-50% of the time ☐₃ Satisfied with it 75-90% of the time

☐₂ Satisfied with it 50-75% of the time ☐₄ Satisfied with it 90-100% of the time

D. Please check one box to indicate your future plans:

☐₁ Plan to continue as life underwriter ☐₃ Hope to get Home Office position

☐₂ Future plans undetermined at this point ☐₄ Hope to become Supervisor or Manager of an Agency

E. With the same amount of effort do you think you would make more or less money in some other line of work than you are making in your present job?

☐₁ More ☐₂ Less ☐₃ About the same

F. Do you think you would get more or less satisfaction from doing some other job than you get from your present one?

☐₁ More ☐₂ Less ☐₃ About the same

(Please turn to next page)

FIG. 3. Third page of Job Satisfaction Questionnaire.

YOUR OWN SUGGESTIONS

Here is your chance to tell how management can do a better job. Take a little time to say what you really think.

G. What could your Manager or General Agent do to improve your job and to help you become more successful?

H. What could your Home Office do to improve your job and to help you become more successful?

Please check the appropriate box under each heading below:

Age last birthday	☐₁ Under 40	☐₂ 40 or over
Number years as a full-time life underwriter	☐₁ Under 2 years	☐₂ 2 years or over
Production last year (Ordinary only)	☐₁ Under $200,000	☐₂ $200,000 or over

Company..... *Agency*.....

Please mail promptly in prepared envelope. Thank you!

Printed in U. S. A. 1946.
Copyrighted, L.I.A.M.A.

FIG. 4. Fourth page of Job Satisfaction Questionnaire.

attitude question was to obtain a check for the answers to the specific questions. It will be noted that the agents tended to reply somewhat more favorably to the general job question than they did to the group of specific questions.

Records are available for 5,758 agents representing the twelve largest companies surveyed. Table 1 shows how their votes were distributed by satisfaction categories.

Table 1

Distribution of "Votes" by 5,758 Agents from 12 Large Companies

On 43 Specific Job Questions			On One General Job Question	
N	%		*N*	%
105,792	43	Extremely Satisfied	2,702	46
90,254	36	Satisfied	2,224	39
36,045	15	Balanced Feelings	696	12
10,713	4	Dissatisfied	107	2
4,790	2	Extremely Dissatisfied	29	1
247,594	100		5,758	100

Thus, when the agents were asked directly how they liked their jobs, 85% answered "Satisfied" or "Extremely Satisfied." Asked how they liked a number of specific things about their jobs, the per cent dropped to 79%. The latter index is regarded as the more accurate one. The fact that the former one is higher suggests that the several items in the questionnaire are not ones to produce an unduly rosy picture of agent attitudes. In no company has the score on the single job item been lower than the composite from the 43 items, although in one company the two per cents were the same.

Job Items Rated Highest and Lowest. Based on the per cent of satisfied votes to the total votes cast, the following ten items (roughly the top quarter) were rated highest.

Note the high scores given interpersonal relations: half the top-ranking items fit this category. It is significant that none of the ten items above relate to training or to compensation although there were a number of items in the questionnaire devoted to these subjects.

The results from the ten items ranking lowest are shown in Table 3. Here the concentration is on money matters, public relations, and distasteful job duties. Each gets three low scores. The tenth item scoring low relates to initial training.

There is close agreement of scores among the ten companies on the highest-ranking items (Table 2) where the average difference between

Table 2

Highest-Ranking Items (10 Companies)

	Per Cent Satisfied		
	Median Company	Lowest Company	Highest Company
How do you feel about . . .			
1. Your Company's general reputation and public standing?	99	93	100
2. The way life insurance performs its basic functions?	98	93	99
3. The opportunities in your work for performing useful service?	96	94	98
4. The opportunities in your work for interesting experiences?	95	93	97
5. Servicing your own policyholders?	92	88	96
6. Your Company's underwriting practices?	91	90	92
7. The cashier in your agency?	90	88	94
8. Your agency's location and work surroundings?	89	87	95
9. Your Manager?	88	82	91
10. Fairness and impartiality your Manager shows in his dealings with his agents?	88	81	91

Table 3

Lowest-Ranking Items (10 Companies)

	Per Cent Satisfied		
	Median Company	Lowest Company	Highest Company
How do you feel about . . .			
35. Prospecting?	71	65	81
36. Preparing briefs and doing other office jobs?	70	59	80
37. Your Company's retirement plan?	69	44	87
38. Servicing orphan policyholders?	68	58	80
39. The public relations programs of life insurance companies?	66	47	74
40. Training you received when you entered the business?	65	55	79
41. The regularity of your income?	59	42	82
42. Your earnings so far this year?	50	44	72
43. Underwriters of other life insurance companies?	49	30	64
44. The attitude of the public toward life insurance agents?	45	23	55

extreme scores is 6.4%-points, and the greatest difference is only 10%-points. Consider the situation with the lowest-ranking items. Here the average difference between extreme scores is 28.7%-points and the greatest difference is 43%-points! This difference occurs on the question asking the agent his attitude toward his company's retirement plan. Some insurance companies have made more liberal provisions than others for retirement, and naturally the scores reflect company differences. Similarly on the question about regularity of income which stands 41st on the list of 44. While almost all established insurance agents are compensated on a strictly commission basis and experience "peaks"and "valleys" in earnings from month to month, some companies have worked out plans for leveling the income of their individual agents.

Important Questions. The items standing highest and those standing lowest are not necessarily the most *important* items. Almost all agents (93% to 100%) are satisfied with their company's general reputation and public standing (see Table 2), but when the scores on this item are correlated with a satisfaction index based on the questionnaire as a whole it develops that the item is not important to job satisfaction. (rho = −. 04). The ten items, in order, showing the *highest* correlations with the satisfaction index are:

	Rho and PE
How do you feel about . . .	
Your company's cooperation with your agency in helping underwriters?	.88 ± .05
The recognition your Manager gives your ability and effort?	.84 ± .06
Your company's interest in the underwriter as an individual?	.77 ± .09
The opportunity in your work for performing useful service?	.72 ± .10
The other underwriters in your office?	.70 ± .10
Your company's desire and willingness to make improvements in agency matters?	.67 ± .11
The opportunities open to you for professional growth?	.64 ± .12
Interviewing prospects?	.64 ± .12
The attitude of your family toward your selling life insurance?	.62 ± .13
Your company's sales promotional materials?	.61 ± .13

These items are the ones which appear to affect the job attitudes of life insurance agents most. Four of them relate to company matters. The importance of family and agency relationships is evident, too. Conspicuous by their absence from this list are the compensation items in the questionnaire.

Nine of the 44 items show a negative correlation with the satisfaction index. All are insignificant with the possible exception of one.

This item reads, "How do you feel about the attitude of the public toward life insurance agents?" Here rho = $-.44 \pm .17$. The relationship certainly is not a close one, but there seems to be some tendency for the satisfied agents to feel less than pleased with the attitude of the public.

Scores by Age, Experience, and Production. Each agent was asked to supply several items of personal information—age, experience, and production. A company officer, seeing this information, might be able to identify the agent in some cases, but only composite figures were reported to the companies and of course no one in the Management Association knew the individuals sending in the blanks.

The personal data were needed to analyze satisfaction scores by different groupings of agents. The results of an analysis of several thousand cases are presented in Table 4.

Table 4

Comparison of Satisfaction Scores by Age, Experience, and Production (8 Companies)

	Satisfaction Score*
Age, under 40	4.19
Age, 40 or over	4.12
Experience, under 2 years	4.29
Experience, 2 years or over	4.10
Production, under $200,000	4.15
Production, $200,000 or over	4.11

* Maximum score = 5.00 (*Extremely Satisfied*); minimum score = 1.00 (*Extremely Dissatisfied*). The *Satisfied* score is 4.00, so these scores, based on the agent's actual "votes," show a level of satisfaction between *Extremely Satisfied* and *Satisfied* but much closer to the latter.

Of the three comparisons shown in Table 4, only the second one is significant,—the neophyte agent expresses more job satisfaction than the veteran agent. On such matters as initial training, which undoubtedly has improved greatly in recent years, the differences in the scores are especially dramatic.

Although there are reliable differences on several compensation items between the small and large producers, the closeness of the over-all satisfaction scores is noteworthy. Certainly high earnings alone do not guarantee high job satisfaction among life insurance agents!

Satisfaction scores were lowest in metropolitan centers, highest in small towns and rural areas. Women agents scored considerably higher than men. Formerly there was a space on the questionnaire to indicate

sex. There is a substantial negative correlation between company size and satisfaction score (rho = $-$ .76 $\pm$ .07 for 16 companies).

Additional Findings. Following the main body of the questionnaire were eight special sections. The first six were alternate-choice questions and the last two were fill-in questions. Some of the most interesting data in the surveys are contained in the answers to the special sections.

Training. Asked, "Do you feel the need for more training?" 78% replied Yes and 22%, No. In no company did the affirmative vote fall below 70%. Life insurance agents like to regard their work as of professional level.

Supervision. The word "supervision," unfortunately, has come to mean close checking on petty matters to some people. Many life insurance agents regard themselves as independent businessmen. They cherish their freedom of operation. We were careful, therefore, not to use the word "supervision" in the question asking about supervision. The wording of the second special question was, "Do you feel the need for closer direction of your activities?" The vote was 45% for and 55% against closer direction. The fact that almost half the agents ask for more "supervision" comes as a surprise to many insurance executives, for they too have cultivated feelings of independence among their field men.

Per Cent Time Satisfied. The third special section serves as a check against the earlier data. Recognizing that everyone experiences ups and downs in his job, this section attemps to get at the typical job attitude. It was decided arbitrarily that no one could be called satisfied unless he checked one of the last two boxes, indicating satisfaction with the job, on the average, at least three days in four. Eighty-two per cent of the agents checked these boxes. It will be noted that this per cent is only slightly greater than the 79%-index determined from the main body of the questionnaire.

Future Plans. If a man plans to continue in his present job or at least in his present field of work, it may be concluded that he enjoys a fair measure of job satisfaction. On the other hand, if he is uncertain about his future plans, it is reasonable to conclude that he may be dissatisfied with his present occupation. In phrasing this question efforts were made to give the dissatisfied an "out" so that the particular choice would not be disturbing. Only 12% indicated, "Future plans undetermined". Fifteen per cent aspired to promotions[5] and the remaining 73% preferred to continue as agents. If the 12% represents the dissatisfaction index, it appears low.

[5] Such "promotions" often involve earning *losses*, particularly during periods of prosperity when insurance sales are relatively easy.

Earning Possibilities in Other Jobs. It was noted earlier that compensation matters did not appear on lists of top-ranking or important items related to job satisfaction. Asked about current earnings, which incidentally, are at a peak, only 50% voted 4 (*Satisfied*) or 5 (*Extremely Satisfied*).

An "earnings question" was inserted in the special section. It read, "With the same amount of effort do you think you would make more or less money in some other line of work than you are making in your present job?" Three choices were offered: *More, Less, About the Same.* The per cents checking these were 16%, 52%, and 32%, respectively. If we split the last per cent between the first two, we can generalize by saying that one agent in three feels he can do at least as well financially in some other line of work while two in three feel selling life insurance is as good a bet as they can find among the gainful occupations.

It should be pointed put that there is no necessary discrepancy between these figures and the earlier ones. Insurance agents, like psychologists and others, may feel some dissatisfaction with current earnings and at the same time recognize that they might not earn more in another line of work. Consider an agent earning $3,000. He may feel this is less than he needs, and he may feel that he is worth $4,000 a year to his company. But he may tell you candidly that he never made more than $40 a week before entering the life insurance business, and he may add that he doubts if he could equal his present earnings of $3,000 in any other line of work.

Again, it is conceded that a satisfied attitude may not always be the most desired attitude. We like the slogan of the milk company which advertises: "Our cows are not contented; they are always trying to do better."

Satisfactions in Other Jobs. This question was similar to the last one. It asked, "Do you think you would get more or less satisfaction from doing some other job than you get from your present job?" Here the vote was 6% for *More*; 76% for *Less*; and 18% for *About the same.*

In every survey and at several points in each survey it is noted that insurance men speak with enthusiasm of the deep satisfactions of their work. Only two of the 44 items in the main body of the questionnaire ranked above "Opportunities in your work for performing useful service." There undoubtedly is a connection between this and the vote by 94% of the agents that selling life insurance provides as much satisfaction for them or more satisfaction than any other occupation.

Agents' Suggestions to Their Managers. In this section and in the last one the agent was offered an opportunity to tell what management —agency management and company management—could do to improve

his job and to help him to become more successful. Space was provided for writing approximately 100 words in each section. Some agents found this space inadequate and continued their remarks on the margin of the questionnaire or on separate paper. A large number of agents commented simply: "Everything possible is being done"; "Well satisfied"; "Nothing to suggest"; etc. The average agent made two or three suggestions.

The comments and suggestions are typed and passed on verbatim in a company report. A grouping by topics of the most-frequently-made comments to Managers from eight companies shows the following ten standing at the top of the list:

Keep in closer touch with us.
Provide us with better office help and working conditions.
Give us better training and supervision.
Build up our morale.
Treat us fairly and impartially.
Have more and better agency meetings.
Help us to do a better job of prospecting.
Get to know our problems better by doing more practical work.
Reduce your personal business and devote more time to us.
Train us more in the technical aspects of our work.

Easily the leading suggestion, expressed in one way or another, is for a closer and more understanding relationship between Manager and agent. There is nothing of a revolutionary nature in any of the suggestions—nothing, in fact, that an alert Manager would not *wish* to do to promote the general welfare of his associates as well as the sales in his agency.

Agents' Suggestions to Their Companies. It should be understood that these suggestions are addressed primarily to the agency departments of the companies. Of course the agency department operates in conjunction with other departments—actuarial, legal, investment, policy issue, underwriting, medical, etc.—as well as in accord with the decisions of top management. Usually a job satisfaction survey is not begun until it has the full approval of the president and other company officers. If the policy-making officers are interested in the idea of a survey from the beginning, they are, of course, more inclined to give attention to the findings at the end. Parenthetically, it seems that the higher one goes in a company, the more acceptance he finds for studies of this sort: sometimes lesser officials are fearful of upsetting the apple cart.[6]

The most-frequently-made suggestions, in order, addressed to the companies were:

[6] For a discussion of this problem see McMurry, R. N. Management's reactions to employee opinion polls. *J. appl. Psychol.*, 1946, **30**, 212–219.

Revise our compensation plan.
Improve our retirement plan.
Get closer to the field—more field trips, more talks with agents, etc.
Advertise more.
Liberalize underwriting practices.
Furnish us with more useful sales tools.
Reduce red tape—speed up policy issue, etc.
Issue more competitive policy lines.
Lower net costs.
Pay us for servicing old and orphan policyholders.[7]

These suggestions, too, are reasonable ones, and most companies are glad to meet their agents half way.

Follow-Up. Speaking at a large gathering of life insurance men, a prominent company officer had this to say regarding the follow-up of the job satisfaction survey which had been made a short time earlier among his 3,500 agents: "It's like a hand grenade,—once you pull the pin you have to do something with it or it may destroy morale rather than build it."

The Agency Management Association takes the position that it is primarily the company's responsibility to do something with the survey findings. That is not to say that the Association maintains a wholly disinterested attitude. It is impossible, having made a survey of a company, not to become interested in its operations. Also, someone present when the survey report is presented and discussed is certain to ask: "What do we do now? What are the best ways to make an effective follow-up of our report?"

Based on the experiences of the first companies making surveys, the Association developed a list of ten follow-up suggestions. This list is made available to companies completing surveys. Of course some companies follow up their surveys more vigorously than others, but all seem to attach considerable importance to the findings and to have a sincere desire to make adjustments and changes in their procedures in accordance with the suggestions of their agents.

Repeat Surveys. The last suggestion on the follow-up sheet reads, "Repeat surveys should be planned every 3 to 5 years to give a moving picture of the company's progress from the agent's point of view." This item is emphasized in personal contact between the Association and company officers. The importance of determining in what direction the company is moving and at what speed is stressed. Attention also is given the thought that the company's primary competition is with itself rather than with the four hundred other life insurance companies operating in the United States. Although in no instance has the recom-

[7] In other words, pay us for doing work which we feel should be done but for which no fee or commission is provided now.

mended 3 to 5 years elapsed since the initial survey, three companies already have requested repeat studies.

Summary

Believing that life insurance agents have attitudes and opinions which are important, the Agency Management Association embarked upon a series of job satisfaction surveys in October, 1943. Seventeen company surveys, involving 6,992 agents have been made and others are scheduled. The surveys are useful in helping the companies surveyed to assess their strengths and weaknesses, and they have values to the life insurance business as a whole. The response to, mechanics, and results of the surveys have been discussed. In general, the morale of insurance agents seems good, the satisfaction index comparing favorably with those reported in earlier attitude studies of workers in other occupations.

Received April 10, 1946.

The Psychological Barometer of Public Attitudes *

Henry C. Link

The Psychological Corporation, New York, New York

The Psychological Barometer is the oldest nation-wide poll of public opinion and buying habits now in existence. Begun in September 1932, these surveys are made four times a year with 10,000 personal interviews and twice a year with 5,000 interviews.

The present survey was made with 5,000 interviews, during October 1946, by 460 interviewers under the supervision of 121 psychologists and in 124 cities and towns. It represents a true cross-section of the urban population. Two questionnaires were used, one with one-half the sample, or 2,500 people, the other questionnaire with the other half of the 5,000 people. These two sub-samples were closely comparable by geographic, sex, socio-economic, and other criteria. Each question reported in this study was made with one or the other of these 2,500 samples. About 31% of the total sample were union members or families with a union member. Further details of the survey are given at the end.

What Does Americanism Mean to the American People?

What is Free Enterprise? What is Capitalism? What is Socialism? What is Democracy? Fascism? Communism? What is the American Way? *What is Americanism?*

This, in short, was the logic and the controversial background leading up to the present survey. In a series of fifteen test or pilot studies, including a test made with fifty professors of American history in as many universities throughout the country, about thirty questions on certain phases of American life were developed. The answers of urban Americans to the first seven of these questions are reported herewith.

Are Unions and the Closed Shop Good Americanism?

The unions and the closed union shop have come to represent a new way of life in America. Our April survey revealed that 67% of the urban public would vote for Congressional candidates who "promised to support stronger laws to regulate the unions," whereas only 20% said they would vote against such candidates. Even among union members, 60%,

* This article is a "prior publication," the author paying complete costs. The scheduled 80 pages per issue is thereby increased by the corresponding amount; thus the "early publication" of this article is a direct contribution to the subscribers of the *Journal of Applied Psychology* without handicap to those authors whose articles are accepted and printed in their regular turn.

as compared with 69% among non-union members, said they would vote for candidates who promised to support such laws. In the present survey the overwhelming majority of urban Americans believe that the right to organize into unions is good Americanism, as indicated by their answers to this question:

Q. "Is it good Americanism for workers to be allowed to organize into unions?"

Answers	Total	Union Members	Non-Union
Yes, good Americanism	82.4%	89%	80%
No, not good Americanism	9.3	5	11
Don't know	8.3	6	9
Total Interviews	2,500	795	1,705

Immediately following this question came the question with regard to the closed union shop. It will be remembered that three states (Nebraska, Arizona and South Dakota) in the recent election voted to outlaw the closed union shop by large majorities.

Q. "Is the closed union shop where every worker must be a union member before he can get or hold a job good Americanism or bad Americanism?"

Answers	Total	Union Members	Non-Union
Bad Americanism	67.3%	55%	73%
Good Americanism	21.3	34	15
Don't know	11.4	11	12

The interesting point here is that even among union members a majority considers the closed union shop bad Americanism. There were important differences by the socio-economic groups, as might well be expected. In the following table the "A" group represents the executive and ownership group; the "B" group primarily the white-collar and professional and semi-professional group; the "C" group the large skilled and semi-skilled industrial and factory workers, including transportation workers, probably the group in which unions are most numerous; the "D" group the less skilled workers and laborers.

	Socio-Economic Groups			
Consider the Closed Shop:	A	B	C	D
Bad Americanism	82%	76%	64%	54%
Good Americanism	10	15	23	31
Don't know	8	9	13	15
Total Interviews	250	750	1,000	500

Should All Races Have Equal Job Opportunities?

Whatever the arguments about the *American Creed* and the *American Way*, the tenet, *equal opportunity for all*, has been considered fundamental. And yet, this is the very point at which racial differences, especially with respect to Negroes (Jews are not considered a race), have retarded the realization of this ideal. What do Americans believe?

Q. "Jews and Gentiles, Negroes and whites, all should have an equal chance at any job: is that good Americanism or not?"

Answers	Total	Socio-Economic Groups			
		A	B	C	D
Good Americanism	84.1%	85%	84%	83%	86%
Bad Americanism	11.6	12	12	13	8
Don't know	4.3	3	4	4	6

The above indicates a great degree of unanimity among people in all walks of life on this basic issue. Even by geographic areas the belief is much the same although not so strong in the South.

Answers	East	Mid-West	South	Far West
Good Americanism	90%	85%	69%	88%
Bad Americanism	6	11	24	9
Don't know	4	4	7	3
Total Interviews	925	788	515	272

The issue of racial equality is brought into still sharper focus by this question, asked of the same people:

Q. "The belief that the Negroes are inferior to the whites: is that good Americanism or bad?"

Answers	Total	Socio-Economic Groups			
		A	B	C	D
Bad Americanism	72.2%	72%	72%	74%	69%
Good Americanism	17.5	16	19	16	17
Don't know	10.3	12	9	10	14

Here again, people of all socio-economic groups think pretty much alike. By geographic sections, however, there are differences, with a substantial difference in the South.

Answers	East	Mid-West	South	Far West
Bad Americanism	77%	74%	59%	77%
Good Americanism	13	15	30	14
Don't know	10	11	11	9

The $25,000 Limitation on Salaries

In the modern trend toward totalitarianism and socialism, one of the common proposals has been the limitation of all salaries to $25,000 a year. This objective has already been closely approached in the United States through the operation of the personal income tax. In view of the basic nature of this issue, the following question was asked:

Q. "No one should be allowed to make or keep more money than $25,000 a year; is that good Americanism or bad?"

Answers	Total	Socio-Economic Groups			
		A	B	C	D
Bad Americanism	65.4%	81%	76%	61%	50%
Good Americanism	26.5	14	19	30	38
Don't know	8.1	5	5	9	12

Even in the large class of skilled and semi-skilled workers, constituting about 40% of the population, twice as many consider this limitation of salaries bad Americanism as consider it good Americanism. Among union members alone, 60% considered this limitation on salaries bad Americanism as compared with 32% who considered it good.

The Individual VS. The State

Probably the basic issue in the world today is that between individualism and socialism, between self-reliance and reliance on an all-powerful state, between the responsibility of the individual and the local community as contrasted with a concentration of power in a highly centralized government, between individuals who want as little government as possible and individuals who, for the sake of social security from the cradle to the grave, are ready to accept the intensive government controls which must go with it. To frame a question which went to the heart of this fundamental issue was extremely difficult, if not impossible. This was the question finally used:

Q. "Which do you think is better Americanism: (a) every man should accept the responsibility for getting his own job and a living, or (b) the government should see to it that every man has a job and a living?"

Answers	Total	Socio-Economic Groups A	B	C	D
Every man should accept responsibility	75.8%	88%	85%	75%	58%
The government should see to it	18.2	8	11	19	32
Don't know	6.0	4	4	6	10

Even the "D" socio-economic group, representing the 20% of the population at the lower end of the economic and cultural scale, considers personal responsibility as better Americanism by a ratio of almost two to one. Among union members, 68% consider personal responsibility as better Americanism, whereas the figure is 79% among non-union members. This is remarkable because, theoretically, if not actually, union members must surrender a good part of their personal independence and self-reliance to the over-all plan of a dominant union.

Is Private Capitalism Good Americanism?

The question of *private capitalism* versus *socialism* or *state capitalism* was put to our sample of 5,000 people in two different forms. In one questionnaire, used with 2,500 people, the question was put in the direct terms of private capitalism and the profit system, as follows:

Q. "Is private capitalism and the profit system good Americanism or bad?"

Answers	Total	Socio-Economic Groups A	B	C	D
Good Americanism	67.4%	85%	76%	63%	55%
Bad Americanism	15.4	9	13	17	18
Don't know	17.2	6	11	20	27

In every socio-economic group, even the "D" group, a substantial majority considers private capitalism as good Americanism. The ratio for the group at large is more than four to one. Union members as contrasted with non-union members think as follows:

Answers	Union Members	Non-Union
Good Americanism	59%	72%
Bad Americanism	20	13
Don't know	21	15

In testing our questions it was obvious that quite a few people do not clearly understand the meaning of private capitalism and the profit system. This accounts in part for the large number of "don't know" answers.

Government Ownership of Large Businesses

To the other half of our sample, or another 2,500 people, the issue of private capitalism versus socialism was put in these terms:

Q. "The Government should own and run large businesses such as the railroads, telephone and telegraph, life insurance, gasoline companies, etc.; would this be good Americanism or bad?"

Answers	Total	Socio-Economic Groups			
		A	B	C	D
Bad Americanism	68.6%	85%	78%	64%	55%
Good Americanism	21.8	8	15	26	30
Don't know	9.6	7	7	10	15

Aside from the lower frequency of "don't know" answers, the results here are very much like those to the preceding question on private capitalism and the profit system. Even in the "C" and "D" groups, about twice as many consider Government ownership bad Americanism as consider it good Americanism. Union members and non-union members answered this question as follows:

Government Ownership is:	Union Members	Non-Union
Bad Americanism	63%	71%
Good Americanism	28	19
Don't know	9	10

The Public Reverses Itself on the OPA

One of the fastest and most sweeping reversals of public opinion we have encountered since our first poll in 1932 is that toward the OPA. Last spring several polling organizations showed substantial majorities in favor of the OPA. Our own survey, in April, using the question: "Do you think that the powers of Chester Bowles and the OPA should be increased or decreased?" showed that 54% of the people thought the

powers should be increased or left as they were, whereas 32% thought that they should be decreased. In the present study we asked the question as follows:

Q. "What do you think would do most to keep prices down: the OPA and its price ceilings; or free competition by business without any OPA?"

Answers	Total	Socio-Economic Groups			
		A	B	C	D
Free competition	65.1%	71%	69%	65%	56%
OPA and price ceilings	26.1	20	25	25	33
Don't know	8.8	9	6	10	11

These interviews were made during the first two weeks in October and indicate that President Truman correctly estimated the public temper in respect to meat controls and other OPA controls by his rapid suspension of these controls during the latter half of October. The answers by socio-economic groups indicate that the rejection of the OPA in favor of free competition was quite general in all socio-economic groups.

The Shortage of Houses, Automobiles, Refrigerators, Etc.

The attitude toward the OPA and other government agencies in the control of materials and prices was reflected by the answers to two questions, asked in opposite forms of the questionnaire:

Q. "Which is most to blame for the shortage of houses and building materials: the labor unions, the manufacturers, the OPA and other government bureaus, or others?"

Answers	Total	Union Members	Non-Union
OPA and other government bureaus	43.1%	47%	42%
Labor unions	18.7	11	22
Manufacturers	11.8	14	11
Others	10.5	10	11
"All three groups"	4.0	3	4
Don't know	18.4	20	17

(*Note:* The total per cents in this and the next question add up to more than 100 because some people gave more than one agency as their answer.)

Q. "Which is most to blame for the shortage of automobiles, refrigerators and other goods like these: the labor unions, the manufacturers, the OPA and other government bureaus, or others?"

Answers	Total	Union Members	Non-Union
OPA and other government bureaus	37.2%	42%	35%
Labor unions	32.3	23	36
Manufacturers	14.0	18	13
Others	8.8	10	9
"All three groups"	6.1	5	6
Don't know	12.2	10	13

The Standard of Living Under Inflationary Conditions

All available Government statistics show that factory and industrial take-home pay since 1941 has risen considerably more than has the cost of living. However, the psychological reaction to this situation is another matter. We have an excellent example of this in connection with questions about electric light and power rates. Our surveys have shown that a large proportion of people believe that the rates for electricity have gone up in recent years, which is contrary to the facts. What has gone up is the use of electricity through the addition of radios, electric refrigerators, more and larger electric light bulbs, vacuum sweepers, etc. The total bill has gone up, but not the prices. Therefore, what people think of their living standards under present conditions is even more interesting and possibly even more important than the actual facts. Since October 1941 we have been asking the following question:

Q. "Is your family more prosperous (or better off) today than two years ago, less prosperous, or the same?"

Answers	Oct. 1941	Oct. 1943	Oct. 1945	Apr. 1946	Oct. 1946
Better off	38%	29%	32%	26%	31%
The same	47	46	51	48	44
Worse off	15	23	15	24	22
Uncertain	—	2	2	2	3

As compared with the Barometer in April 1946, the October Barometer shows no decline in people's living standards. More people (31%) consider themselves better off today than consider themselves worse off (22%), whereas in the April 1946 Barometer, 24% considered themselves worse off. There have been price increases as well as price de-

creases since the middle of October which this survey does not reflect. Also, these answers reflect not only the cost of living but the extent to which people are employed and earning money. A sharp drop in employment, even if accompanied by a sharp drop in the price level, would quickly reflect itself in the answer to the above question. The results by socio-economic groups show that the "B" group, the white collar workers who have been heralded as the chief victims of inflation, consider themselves about as well off as do the "D" and "C" industrial and wage earning groups:

	Socio-Economic Groups			
Answers	A	B	C	D
Better off	28%	32%	32%	29%
The same	50	44	42	45
Worse off	20	21	23	21
Uncertain	2	3	3	5

The unions are supposed to give their members a great advantage in obtaining higher living standards. Here are the results by union and non-union membership:

Answers	Union Members	Non-Union
Better off	33%	30%
The same	42	45
Worse off	22	21
Uncertain	3	4

Interestingly enough, the answers of union members indicate that they considered themselves no better or worse off than did the non-union members.

The Diminishing Prospects for Peace

Before the United Nations began its recent session, the city of New York organized a demonstration for them in the form of a parade and reception. According to the newspaper reports describing the occasion, the crowds lining the streets as the representatives of the nations went by manifested little enthusiasm. In fact, their behavior showed almost complete apathy. If these reports were correct, they do no more than reflect the growing opinion of the American public

toward the prospect of peace. In six of these surveys we have now asked the following question:

Q. "After this war (or, now that the war is over) do you think that we will make a peace settlement that will last or do you think that we will have another world war in twenty-five years or so?"

Answers	Feb. 1943	Oct. 1944	Apr. 1945	Oct. 1945	Apr. 1946	Oct. 1946
Will have another war	43%	54%	51%	59%	62%	74%
Will make a lasting peace	47	28	33	28	24	18
Don't know	10	18	16	13	14	8

This represents the largest increase yet in the percent of those who expect another war, from 62% to 74%. The expectation of Russia as our next enemy shows a large increase, whereas the belief of another war with Germany and Japan has almost disappeared.

Q. "Who do you think will be our next enemy?"
Answers by Those Who Said There Would Be Another War:

Country Named	Oct. 1944	Apr. 1945	Oct. 1945	Apr. 1946	Oct. 1946
Russia	29%	27%	37%	45%	56%
Germany	9	6	2	2	2
Japan	5	3	5	1	1
England	4	4	3	4	3
China	1	1	1	1	1
Don't know	6	10	11	9	11
Total	54	51	59	62	74

Explanation of the Survey

All interviews were made in the home, but only one in a family; half were made with women, half with men. The 5,000 interviews were distributed by four socio-economic groups referred to in the previous tables as A, B, C, and D. This distribution was made in accordance with the socio-economic maps in each locality according to which the local supervising psychologist assigned the calls to be made by streets and blocks. The great differences between the thinking of these various socio-economic groups are shown in some of the tables. These differences, incidentally, are also an indication of the thoroughness with which these interviews have been distributed by socio-economic levels.

Received December 23, 1946.

Bibliography

The Psychological Corporation. A study of public relations and social attitudes. *J. appl. Psychol.*, 1937, **21**, 589–602.

Link, H. C. A study of public opinion and morale. *J. appl. Psychol.*, 1941, **25**, 636–645.

—— The eighth nation-wide social experimental survey. *J. appl. Psychol.*, 1943, **27**, 1–11.

—— The ninth nation-wide social experimental survey. *J. appl. Psychol.*, 1944, **28**, 1–15.

—— The tenth nation-wide social experimental survey. *J. appl. Psychol.*, 1944, **28**, 363–375.

—— The Psychological Corporation's index of public opinion. *J. appl. Psychol.*, 1945, **29**, 103–107.

—— The Psychological Corporation's index of public opinion. *J. appl. Psychol.*, 1946, **30**, 1–9.

—— The Psychological Corporation's index of public opinion. *J. appl. Psychol.*, 1946, **30**, 297–309.

How Many Foot Candles? *

C.-E. A. Winslow

School of Medicine, Yale University

The problem of artificial illumination is of primary importance in the sanitation of occupied spaces. Adequate light conduces materially to comfort and efficiency in the duties and avocations of daily life; and faulty illumination causes eyestrain which may be accompanied by reflex functional disturbances of many other organs of the body. Sound standards for artificial lighting are therefore essential tools for the sanitarian, whether he is working in the school, the home, the office, or the factory.

The leadership of the Illuminating Engineering Society in preparing standards of lighting for various types of occupied spaces was therefore welcomed by the public health profession. During the past quarter-century, this organization—often in cooperation with the American Institute of Architects, and, usually, under the rules of procedure of the American Standards Association—has issued a large number of codes of Recommended Practice for Lighting.

Those codes have added materially to our knowledge of this subject; and in relation to color of light, distribution of light and brightness contrast they are fundamentally sound. Prior to 1940, they were practical and helpful with regard to intensity of illumination. The code on School Lighting prepared in 1938, for example, called for 15 foot candles for classrooms, shops, and offices, 25 for sewing and drafting rooms and 30 for sight-saving classes. These standards are reasonable and are supported by adequate evidence. Since 1940, however, under the leadership of Matthew Luckiesh of the Nela Park Laboratories at Cleveland, there has been an extraordinary and progressive tendency to increase these standards so that with every report the minimum levels rise; and, beyond the minima set, it is suggested that still more astronomical intensities would really be desirable. Twenty-five to 50 foot candles are specified for average visual tasks, 100 foot candles or more for severe and prolonged visual work, and hundreds and even thousands of foot candles are indicated as desirable in certain cases.

* Editorial by Dr. Winslow in *American Journal of Public Health*, September, 1946, Vol. 36, pp. 1040–1041. Reprinted by permission.

Sanitarians have generally viewed these extravagant standards with justified suspicion. As long as J. E. Ives of the U. S. Public Health Service was alive, we always turned to him for counsel since, though he was a physicist, he understood physiology. Since his death we have turned to the psychologists as our best guides; and we are fortunate in presenting as the Special Review Article[1] of this issue a thoroughgoing analysis of the whole subject by the outstanding leader in this field, M. A. Tinker, of the University of Minnesota.

In this article, Dr. Tinker analyzes the evidence presented by Luckiesh and his associates in support of their high demands. He shows how the physicists have treated light intensity as a purely quantitative factor, almost as if when a certain intensity proved good, more was necessarily better. They have ignored the fact that the human eye is an adaptive instrument and that preference at a given moment is a function of prior adaptation; not of permanent comfort. They have confused threshold visibility, which is chiefly physical, with ease of reading above the threshold, which is quite a different problem. They have wholly omitted a consideration of the glare effect incidental to very high illumination. Above all—and this seems inexcusable for physicists—they have paid no attention at all to the shape of the curve of visibility and of visual performance. This curve (as shown by Ives many years ago for letter-sorting in the U. S. Post Office) rises rapidly with increase in illumination and then levels off to almost a straight line. Luckiesh computed that it would take 30 foot candles to make newspaper text matter equivalent in visibility to 8-point book type under 10 foot candles illumination; on the other hand, Tinker found that no improvement in reading performance with newspaper type actually resulted with increases beyond 7 foot candles. Lythgoe has found that, under certain conditions, visual acuity continues to increase up to 1,000 foot candles; but the inflection of the curve is at 10 foot candles and increase above 20 foot candles is slight. Actually, Tinker concludes that the critical levels of illumination for different visual tasks vary from 3 foot candles for reading large print to 30 for threading a needle; and that to allow an adequate margin of safety, standards for various activities might range from 10 to 50 foot candles; but not above.

The whole problem is strikingly reminiscent of the controversy of some decades ago with respect to school ventilation. In that field, too, the engineers called for standards (30 cubic feet per minute) which were shown to be excessive by physiological research; and, in view of accumulating evidence, the American Society of Heating and Ventilating Engineers changed its standards to conform to real human needs. The Illu-

[1] Tinker, M. A. Illumination standards. *Amer. J. Public Health*, 1946, 36, 963–973.

minating Engineering Society should exhibit similar readiness to follow the truth. As Tinker points out, satisfactory results can only be attained by "coordinating the work of engineers, physiologists, and psychologists." Until such coordination is effected, we shall do well to follow the psychologists, since it is human health and comfort which is at stake.

Some Comments on the Army General Classification Test *

Acheson J. Duncan

Department of Political Economy, The Johns Hopkins University

In a recent article[1] the Personnel Research Section of the Adjutant General's Office presented some very interesting facts regarding the Army General Classification Test (AGCT). The following are a few comments supplementary to that article. They are based on the writer's personal experience with the test in the Army and upon data collected by him while in the service.

What Does the AGCT Measure?

The AGCT is declared to be a test of "general learning ability."[2] That it did very effectively distinguish those of high learning ability from those of low or average ability was obvious to any one who had contact with the test in the service. Care should be exercised in using the test, however, and it is the intention of the following comments to point out in what respect the test may have special biases and weaknesses.

As pointed out by the Staff of the Personnel Research Section, the specifications for the test called for emphasis upon spatial thinking and quantitative reasoning since this was the kind of thinking the Army would have the greatest need for.[3] Accordingly, one third of the test consisted of arithmetic questions and another third of block-counting problems. It would seem therefore that ability to handle material of this kind was given special value by the test. In this connection, it would be interesting to compare the AGCT scores of accountants, engineers, and the like with the general run of scores to see if men who were in the habit of dealing with figures and diagrams did better than average in the AGCT.

Specifications for the AGCT also called for minimizing the time or speed factor.[4] Nevertheless, the large number of questions asked (150)

* Being an economist and not a psychologist, the writer offers these comments as the reactions of a layman to psychological testing. It is hoped that they will be of interest and possibly some stimulus to the trained technician.

[1] Staff, Personnel Research Section, Classification and Replacement Branch, The Adjutant General's Office, "The Army General Classification Test," *Psychological Bulletin*, Vol. 42, No. 10 (December 1945), pp. 760–768.

[2] *Ibid.*, page 760.

[3] *Ibid.*, p. 760.

[4] *Ibid.*, p. 761.

and the time limit of forty minutes gave an advantage to those who worked more rapidly. This was verified to some extent by the negative correlation that the Personnel Research Section found to exist between age and AGCT scores among two selected groups of officers.[5] Additional support is rendered by the analysis of the following section which shows that in groups of men with high AGCT scores there is little correlation between educational level and AGCT score. For those with a higher level of education were usually older men whose reactions were probably slower and who were therefore probably penalized by the speed element in the test.

Although the AGCT was not designed to measure personality traits,[6] the method of scoring the test would appear to have unavoidably given a premium to those who liked to gamble. Four choices were given as answers to each question and the raw score was computed by deducting one-third the number wrong from the number right. On the average, this would give completely blind guessing a raw score of zero. On the test booklet, the man being examined was told that if he did not know the answer he should guess. This implied gambling on partial or imperfect knowledge. Since the man was not told how the test was scored he did not know what was to be gained or lost by guessing and it is likely that the bolder men gained higher scores by taking more chances.

In this same connection it is to be pointed out that the use of separate answer sheets especially hurt the few who had the misfortune to unknowingly skip a question on the test paper. If this mistake was not discovered for several questions, a considerable loss of time was incurred in erasing marks on the answer sheet and in reanswering the questions wrongly marked. Bungling of this kind was thus heavily penalized.

Whether the AGCT distinguished high learning ability from *very* high learning ability is a debatable question. The test itself asked relatively simple questions, and for those with the same level of knowledge and training it might seem that the test distinguished primarily between speeds of reaction. Offhand, therefore, the AGCT might not seem to measure directly a man's ability to solve complicated and abstract problems in which the speed factor is at a minimum. These theoretical or a priori arguments against the AGCT, however, do not appear to be supported by statistical evidence. At the time the Army Specialized Training Program was being organized, men with high school or college education and with high AGCT scores (none less than grade II) were given a special test (OCT-2) which aimed to show their fitness for this Army college training. This was a three hour examination and was a test of higher levels of reasoning power. It asked questions about diffi-

[5] *Ibid.*, p. 765.
[6] *Ibid.*, p. 761.

cult reading material, a complicated time table, a graph, and the like; it also had questions on acquired knowledge of the sciences and elementary mathematics. The writer correlated the scores of 120 men on this test with their AGCT scores and found a product-moment correlation coefficient of .54. This is not a very high correlation, but it would seem to indicate that the AGCT might serve as a rough means of distinguishing higher learning ability.

The AGCT and Education

The Personnel Research Section of the Adjutant General's Office presented in their article several sets of data showing a fair degree of correlation between AGCT scores and extent of formal education. In Table 1 is another set of data pertaining to the AGCT scores and formal

Table 1

AGCT Grades and Formal Schooling of 990 Enlisted Men of a Coast Artillery Regiment

Schooling	AGCT Grades					
	V	IV	III	II	I	Totals
Some Post Graduate work	—	—	—	9	—	9
College complete	—	—	—	9	3	12
College incomplete	—	—	12	30	6	48
High School complete	—	17	103	139	21	208
High School incomplete	—	60	140	85	11	296
Grammar School complete	8	80	64	22	3	177
Grammar School incomplete	32	89	39	8	—	168
Totals	40	246	358	302	44	990

education of 990 coast artillery men. The coefficient of contingency for this table is .54, which agrees very well with the product moment correlation coefficients obtained by the Personnel Research Section.[7]

It is seen from this chart that none who had had some high school training got as low as grade V on the AGCT ,and none who had had some college training got as low as grade IV. Furthermore, college graduates did not get below a grade of II. The relationship stops here, however. For the distribution of those with an AGCT grade of I is roughly the same

[7] In computing the coefficient of contingency, all those with more than a high school education had necessarily to be thrown into one group. For the number of classes so obtained, the maximum value of the coefficient of contingency is .894. (See G. U. Yule and M. G. Kendall, *An introduction to the theory of statistics*, p. 69.) A contingency coefficient of .54 has therefore probably greater significance than a Pearsonian coefficient of the same value.

as the distribution of those with an AGCT grade of II. It is true that none with a grade of I had failed to graduate from grammar school, but on the other hand none who had taken college post graduate work got a grade of I.[8]

The similarity of the distribution of men in grade I with the distribution of men in grade II is attested by a χ^2 of only .91. This indicates that two such sample distributions from the same population might differ to a greater degree than this by chance more than fifty times out of one hundred.[9] Although the numerical differences are thus small, the distribution of men in grade I does show a slightly smaller percentage of men below the modal class than does the distribution of men in grade II.[10] On the whole, it would appear that at high levels for both variables there is some correlation between AGCT score and extent of formal education, but that it is not very great.

An interesting side observation is that 29 per cent of the men had no regular college training but nevertheless got grades of I and II on the AGCT. This suggests that whereas only 7 per cent of the men had had some college training, an additional 29 per cent might have the ability to undertake college work. Even if these figures are only roughly correct, they are a strong argument for Government scholarships to young men who have ability, but lack financial means. It is hoped that the post-war experience of the colleges with returning veterans will be studied with the object of determining the desirability of such government grants.

Retests and Reliability of the AGCT

In the unit in which the writer served, more than 100 men took the AGCT a second time. The interval between tests was at least several months. In 56 of these retests, the men were given comparable tests, i.e. a *1-a* or a *1-b* test followed by a *1-b* or a *1-a* test, or a *1-c* or *1-d* test followed by a *1-d* or a *1-c* test.[11] For this group of retests the average increase in standard score was 11.25 points, or about one half the standard deviation of the tests. Of the total only one man decreased his score, ten men increased their scores from 0 to 4 points, nine from 5 to 9 points,

[8] It was suggested above that this might be explained by the age and slower reactions of these men. It is also possible that the coast artillery did not get its proportionate share of men who had both a high level of education and a high AGCT score. Against this second explanation, however, was the general policy of the Army to assign an equal share of all classes of men to every branch of the service.

[9] See any χ^2 table for $n = 2$. To make this comparison, this distribution had to be telescoped at both ends, making only three cells in each column.

[10] For grade I, 32 per cent of the men are below the modal group; for grade II, 38 per cent. For grade I, 20 per cent are above the modal group; for grade II, 16 per cent.

[11] See Personnel Research Section, *op. cit.*, pp. 762–764, for a discussion of the alternate forms of the AGCT.

sixteen from 10 to 14 points, fourteen from 15 to 19 points, four from 20 to 24 points, and two from 25 to 29 points. The standard deviation of the change in score was 6.77 points.

The average increase of 11.25 points has several explanations. First it must be recognized that retests were supposed to be given to men only when there was some positive reason to believe that the man's first score was not a true measure of his ability. In practice, however, most men who wanted retests were allowed to have them. Usually they hoped to raise their scores sufficiently to make the Officer Candidate level or the level required by a special Army course they wanted to take. In a few cases, men who already had high grades were allowed to take a retest to see if they could raise them still higher.

This "special selection" of the men who took retests is not, however, the main explanation of higher scores on this second test. The more important reasons would appear to be the difference in conditions under which the tests were given and a greater familiarity with the testing procedure (a practice effect). The original test was given to most men within a day or so after their arrival at the reception center. Often they lacked sleep, or were recovering from a farewell party, or felt ill from a battery of shots given them just prior to the test, or just didn't care.[12] One man who increased his score by a very considerable margin said that at the reception center the light had been so bad that he couldn't see the paper clearly. At the time the men took the retests, they had their basic training behind them and had some months of regular Army service. They were consequently more adjusted to Army life and this was conducive to better results on the second test.

Familiarity with the testing procedure was the second principal reason for higher grades. The men now knew what they were up against. They might even have practiced up on "sample" tests sold by commercial agencies or they might have learned how the test was scored and taken advantage of this knowledge. That greater familiarity with the test was a factor in raising grades would seem to be demonstrated by the results of the retests of eight men who took the test a *third* time. For this group, the increase in score from the second to the third test ranged from 7 to 17 points, indicating that the average difference between the second and third tests was about the same as the average difference between the first and second test. This suggests that the practice effect may have been the primary reason for the increase in scores on subsequent tests.

[12] In World War I, however, it was shown that recent (i.e., within 24 hours) typhoid inoculations had no significant effect on Army Alpha scores. See *Memoirs of the National Academy of Sciences*, Vol. 15, p. 809.

The correlation between the first and second scores on these 56 retests was .781 which yields a standard error of measurement of almost 4.7 (standard) points.[13] For 46 other scores on retests in which a *1-a* or *1-b* test was followed by a *1-c* or *1-d* test, the correlation between first and second scores was.788 which is in very close agreement with the other result. The coefficients of reliability obtained by the Personnel Research Section from retests and by the Kuder-Richardson method were somewhat larger than this (.82 to .97) and the standard errors of measurement were somewhat less, but the differences were not great.[14]

The AGCT and Rank in the Service

The AGCT score was used considerably in the Army in selecting men for various jobs and for school courses leading to higher positions. It might be expected therefore that there would be some correlation between AGCT score and rank attained in the Army. To investigate this hypothesis, rank (as measured by pay received) was correlated with AGCT score for the enlisted men of both a "line" and a headquarters coast artillery battery. The coefficient of correlation was .31 for the line battery and .20 for the headquarters battery.[15]

The reasons for these rather low correlations are not hard to find. First, as pointed out above, the AGCT was not intended to be a measure of personality traits, and data presented by the Personnel Research Section shows that there was very little correlation between AGCT scores and leadership ratings at the Infantry and Ordnance Officer Candidate Schools.[16] To the extent, therefore, that rank among enlisted men was based on leadership ability, no significant correlation could be expected.

[13] The standard deviations of the two tests were both about ten points, so that $\sigma_s = 10\sqrt{1 - .781} = 4.7$. The means of the two sets of tests were 101.48 and 112.73, respectively.

[14] *Op. cit.*, p. 765. In those cases in which very high coefficients of reliability were obtained, the standard deviations of the tests closely approximated the full deviation of 20 (standard) points, so that the standard error of measurement was still around 4. Thus if $\sigma = 20$, $\sigma_s = 20\sqrt{1 - .96} = 4.0$. The formula that is used here to estimate the standard error of measurement from the coefficient of reliability is based on the assumption that any factor tending to increase the grades on the second test affects every individual alike. When differences have been computed, therefore, for each individual pair of test scores, and the standard deviation computed for these differences, the standard error of measurement may, on the basis of the same assumption, be computed by dividing the standard deviation of differences by $\sqrt{2}$. Thus for the above data, the standard deviation of differences equals 6.77, and $6.77/\sqrt{2} = 4.78$, which is practically the same as obtained by calculating the correlation between the tests.

[15] There were about ninety cases in each group.

[16] The correlations were .12 and .09 respectively. See *op. cit.*, p. 767.

Probably more important was the tendency for rank to vary with seniority, especially for old and established outfits such as were generally found in the coast artillery. Thus, when the writer correlated rank against length of service for both the line and headquarters batteries referred to above, he found correlations of .70 and .59 respectively. It would seem therefore that roughly one third to one half of the variation in rank[17] is to be explained by length of service, with other factors, such as personality, accounting for the remainder. It would be interesting to know whether AGCT scores show a better correlation with rank in outfits newly organized during the war.

Received March 28, 1946.

[17] $(.70)^2 = .49$ and $(.59)^2 = .348$.

The MacQuarrie Test for Mechanical Ability: II. Factor Analysis

Charles H. Goodman

Radio Corporation of America

A preceding paper[1] on the MacQuarrie Mechanical Ability Test described the experimental use of this test for the purpose of determining its selective capacity in hiring assembly operators in a radio manufacturing company. The present paper reports the results of a factor analysis study of the MacQuarrie Test measures, the seven sub-test scores of 329 radio assembly operators were correlated. The intercorrelations obtained provided the correlation matrix necessary for the factor analysis. The factor technique[2] employed in this analysis was the Thurstone Centroid Method. Table 1 shows the correlation matrix of the seven MacQuarrie Sub-Tests.

Table 1

Correlation Matrix of Seven MacQuarrie Sub-Tests

Test	Tracing	Tapping	Dotting	Copying	Location	Blocks	Pursuit
Tapping	.48						
Dotting	.55	.47					
Copying	.44	.31	.34				
Location	.34	.29	.43	.54			
Blocks	.41	.29	.32	.52	.54		
Pursuit	.43	.29	.36	.48	.44	.46	

After the third factor had been extracted an examination of the third factor residuals was made. These third factor residuals are shown in Table 2. It can be seen in this table that the largest positive entry in any of the new columns is +.07. The largest negative entry in any of the new columns is −.07. On the basis of these two values the maximum projection which could have been obtained from the rotation of these axes would have been +.14. Since a projection of +.14 would only account for two per cent of the total variance of the test, there would have been little value in attempting to name the factor.

[1] Goodman, Charles H. The MacQuarrie Test for mechanical ability: I. Selecting radio assembly operators. *J. appl. Psychol.*, 1946, **30**, 586–595.

[2] Guilford, J. P. *Psychometric methods.* McGraw-Hill Book Co., Inc., New York, 1936.

Table 2

Factor Residuals after Extraction of Third Factor

	−1	+2	−3	+4	+5	+6	+7
−1	+.052	+.014	+.003	−.036	+.052	−.025	+.026
+2		+.035	+.017	+.011	+.002	+.010	+.019
−3			+.070	+.028	−.070	+.027	+.010
+4				+.040	−.001	−.018	+.001
+5					+.070	+.005	+.038
+6						+.034	+.012
+7							+.050

It was therefore considered that a sufficient number of factors had been extracted. The original factor loadings which were obtained are shown in Table 3.

Table 3

Original Factor Loadings and Their Communalities

Test	Original Loadings			Communalities
	I	II	III	h^2
Tracing	.694	−.288	−.115	.578
Tapping	.570	−.350	.144	.469
Dotting	.656	−.315	−.087	.537
Copying	.689	.268	.033	.548
Location	.679	.272	.128	.551
Blocks	.668	.290	.060	.534
Pursuit	.638	.156	−.182	.464

In the psychological interpretation of rotated factors, the writer has accepted Thurstone's[3] criterion of not regarding a loading as significant unless the loading is as large as .40. This criterion is used by Thurstone on the basis that a factor cannot be named with any assurance unless it accounts for a third or a fourth of the total variance of a test.

After the factor loadings had been rotated a solution was reached whereby it was felt that simple structure had been obtained. The rotated factor loadings are shown in Table 4.

Column one of Table 4 rearranged in order of the size of the projections obtained presents the following arrangement: Tracing .369; Dotting .338; Pursuit .325; Copying .116; Tapping .105; Blocks .080; and Location .023.

[3] Thurstone, L. L. *Primary mental abilities.* University of Chicago Press, Chicago, Illinois, 1938.

Table 4

Rotated Factors and Communalities

Test	Rotated Factor Loadings			Communalities
	I	II	III	h^2
Tracing	.369	.445	.492	.576
Tapping	.105	.352	.578	.469
Dotting	.338	.405	.509	.537
Copying	.116	.727	.061	.546
Location	.023	.738	.084	.553
Blocks	.080	.725	.032	.533
Pursuit	.325	.595	.067	.464

None of the loadings in column one are as large as the criterion .40 which has been accepted as the criterion for naming a factor. It is interesting, however, to examine this factor in view of the fact that it appears in three of the tests with the remaining tests being essentially negative. Upon examination of these three tests it appeared to the writer that this factor was probably a visual inspection factor. The operation of this visual inspection is necessary in the Tracing Test in order for the subject to align his pencil before passing through the gaps in the lines to prevent him from striking the lines. Visual inspection is called for in the Dotting Test when the subject is required to align his pencil in order to place the dot in the small circle. The Pursuit Test requires visual inspection in order to follow the particular line one is tracing through the maze of lines in which it is entangled.

At first the writer experienced difficulty in determining what the common factor might be which pervaded these three tests. It was only upon the completion of a motion analysis of the MacQuarrie Test that the writer was able to gain insight into the common feature involved in these three tests. While this factor has been named tentatively as a visual inspection factor by the writer, it should be remembered that the loadings of this factor are not large enough to be considered significant.

The projections found in column two of Table 2 rearranged in order of size are as follows: Location .738; Copying .727; Blocks .725; Pursuit .595; Tracing .445; Dotting .405; and Tapping .352.

Six of the seven loadings in this column are significant, with the tests of Location, Copying, and Blocks carrying the heaviest loadings. The tests of Location and Copying call for the transposition of a point or a series of points from one field of reference to a second field of reference which is different in proportion to the first.

Examination of these tests for the common element among them led the writer to unhesitatingly name this factor to be spatial in character.

Thurstone[4] in his factor analysis of a battery of fifty-six tests, used in his battery the three MacQuarrie Sub-Tests of Blocks, Pursuit and Copying. It is worthwhile noting that in the factor isolated by Thurstone as a spatial factor, the tests of Blocks and Pursuit were found to have significant projections. The test of Copying, which was found to have the second highest loading in the writer's spatial factor, was not found to have a significant projection in the Thurstone Study. It is worthwhile pointing out that it seems that this spatial factor appears to permeate all of the MacQuarrie tests, with the exception of the Tapping Test.

The loadings in column three rearranged in order of size are as follows: Tapping .578; Dotting .509; Training .492; Location .084; Pursuit .067; Copying .061; and Blocks .032.

Only the first three tests in this column have loadings of any significance. Study of these three tests leads the writer to believe that the common element is a controlled manual movement. The Tapping Test requires, besides this controlled manual movement, a ballistic movement, which has been found to be a much used movement in the type of industry where this study was made. This test offers some possibilities for further study in the direction of selecting radio assembly operators. The writer has named this factor a controlled manual movement factor.

Harrell[5] in his factor analysis of a group of mechanical ability tests, had as a part of his test battery the seven MacQuarrie Sub-Tests. Harrell found on completion of his analysis that the MacQuarrie tests of Copying, Blocks and Pursuit were loaded on a factor which he named a spatial factor. This has also been found to be the case in this study. However, the remaining four tests of the MacQuarrie were not significantly weighted on the spatial factor. In Harrell's study, the Dotting Test was found to have its heaviest loading on a factor which Harrell calls a manual agility factor. The Dotting Test shows up most significantly in the writer's analysis on the factor called the controlled manual movement factor. The tests of Tapping and Tracing were also found to be significant on the writer's controlled manual movement factor, but this was not the case in Harrell's manual agility factor.

Summary and Conclusions

A factor analysis of the seven MacQuarrie Sub-Tests was made for the purpose of determining what factors were being measured by this test. The analysis disclosed that the MacQuarrie Test was measuring three

[4] Thurstone, L. L. *Op. cit.*

[5] Harrell, W. A factor analysis of mechanical ability tests. *Psychometrika*, Vol. 5, No. 1, March, 1940.

factors. Two of the factors were clearly identified as a space factor and a controlled manual movement factor. The third factor has been tentatively identified as a visual inspection factor. Indentification of this factor is only tentatively made because the factor loadings obtained were not large enough to name this factor with any assurance.

The findings of this analysis appear to be substantiated in some degree by the findings of two other factor analysis studies in which some of the MacQuarrie Sub-Tests were used. There is some discrepancy in the naming of the controlled manual movement factor by the writer and the factor called by Harrell a manual agility factor. This may only be a difference of nomenclature. The writer's basis for naming the factor as he did, was due to the findings of a time and motion study, made on the MacQuarrie Test.

Received April 17, 1946.

Reliability of the Grove Modification of the Kent-Shakow Formboard Series

Ruth C. Wylie

Connecticut College

After having made an experimental study of the Kent-Shakow Formboard Series, Grove modified it[1] in an attempt to correct what he considered to be certain weaknesses of the original. Grove claimed that, while preserving the most important features of the original, his modification rendered the test more homogeneous in content. He demonstrated that it still discriminated adequately at the adult level. The principal advantage of the modification is the fact that it takes on the average less than fifteen minutes to administer, whereas the original requires from thirty to forty-five minutes.

The original Kent-Shakow Formboard Series was included by Pond in a battery of tests used to select toolmaker apprentices at the Scovill Manufacturing Company.[2] At the Cincinnati Employment Center it was found that this formboard series discriminated among occupational groups[3]. Bingham suggests use of this test among others for estimating an aspect of engineering aptitude.[4]

Correlations with other tests, both verbal and performance, reported by Grove (loc. cit.) support the hypothesis that his modification measures an aspect of general intelligence, but that it will probably be found more heavily weighted with what has more recently come to be known as the spatial visualization factor of ability.

That this is an important aspect of mechanical ability is now fairly well accepted. Unpublished data made available to the writer by Grove indicate that the product moment correlation between his modification and the long form of the Minnesota Mechanical Assembly Test was .63 in a population of 200 young adult, native white, male penitentiary inmates. The modified formboard series was also found to discriminate sharply between a carefully selected group of graduate engineers and a

[1] Grove, W. R. Modification of the Kent-Shakow Formboard Series. *J. Psychol.*, 1939, 7, 385–397.

[2] Bingham, W. V. *Aptitudes and aptitude testing.* New York: Harper, 1937, p. 138 f.

[3] Paterson, D. G., Schneidler, G. G., and Williamson, E. G. *Student guidance techniques.* New York: McGraw-Hill, 1938, pp. 229–233.

[4] Bingham, W. V. *Op. cit.*, p. 172 ff.

group of college men matched for ACE scores but lacking in mechanical interests or experience. The median ACE score for both groups on college norms was at about the 95th centile.

Probably because of the technical difficulties involved, satisfactory information on the reliability of many of the performance tests now in use has never been provided. This deficiency exists largely for those performance tests in which intellectual insight and specific memory of the problems presented makes the repetition of the test something quite different from its first administration. This is particularly true of the modified Kent-Shakow Formboard Series. Thus the retest method of estimating test reliability with this test is largely vitiated.

While some of the data referred to above suggest that this test may have significant utility, to establish its validity satisfactorily will require a great deal of research that would be largely wasted if the test is relatively unreliable. It seemed most logical to try to establish reliability before further research was undertaken. The primary purpose of this study was, therefore, to ascertain so far as possible, the reliability of the Grove modification of the Kent-Shakow Formboard Series. Additional data were gathered on the reliability of several other performance tests believed by Grove to measure abilities closely related to those measured by this test.

Procedure

As part of a larger study, four performance tests and two group tests were administered to 352 boys chosen at random from grades 7-12 of the Beaver Falls, Pennsylvania Public Schools.

1. Grove's Modification of the Kent-shakow Formboard Series.
2. Block Designs Test from the Wechsler-Bellevue Scale.
3. Cube Construction Test from the Cornell-Coxe Scale.
4. O'Connor's Wiggly Block Test.
5. Revised Minnesota Paper Formboard, Series AA.
6. California Test of Mental Maturity.

Data from the two group tests indicated that the sample was reasonably similar to high school populations used in standardizing the group tests.

Since, as stated above, the test-retest method is not suitable for measuring the reliability of performance tests involving the perception of a rational problem, it was decided to try the split-half method. This is frequently the preferred method for use with pencil and paper tests, but is not easily applied to most performance tests.

The formboard series consists of four separate subtests of five recesses each. Splitting the test by subtests does not yield two halves of equal

difficulty. Therefore a new method of recording was introduced. The examiner was trained to record accurate, cumulative time and move readings after each recess, while the test was being administered in the standard fashion. By successive subtractions, the number of seconds and the number of moves spent on each recess could be determined later, giving 20 time scores and 20 move scores for each subject. These raw time and move scores were converted into rate scores, a procedure justified by research to be reported in another paper.

The 20 time rate scores were split into odds and evens, yielding a division of the test into approximately equally difficult halves. The 20 move rate scores were also split into odds and evens, again yielding an approximately equal division of the test.

Reliability coefficients for these time and move scores were then computed for each grade and for the pool of grades 9-12. The latter group proved homogeneous enough to be put into this larger pool separate from grades 7-8.

Time scores for each part of the other performance tests used were converted into rates and split-half reliability coefficients computed, although odds versus evens did not yield as nearly equal halves in these tests as in the case of the formboard series. The O'Connor Wiggly Block Test was eliminated because its retest reliability was very low, confirming the coefficients found in the literature for this test. In order to see whether any combination of the three remaining tests would yield a battery more reliable than any one test alone, appropriate calculations were made.

Results

The results are summarized in Table 1.

Discussion

It will be noted that, for the formboard, time scores were consistently more reliable than the move scores. While the magnitude of these differences probably fluctuated to some extent with the chance factors of sampling, the trend seems unequivocal. Combining time and move scores probably does not increase the reliability of the test in so far as the results with the eleventh grade may be accepted as representative.

The results obtained seem to indicate that the Grove modification of the Kent-Shakow Formboard Series is about as reliable as the Block Designs Test and more reliable than the Cube Construction Test. In the latter case, however, it must be noted that the technical difficulties of applying the split-half technique to this test may have lowered the index of reliability more than should have been the case. This is largely

speculative, however. Combining the performance tests used in this study apparently yields a battery of greater reliability than that of any of the three tests taken separately. Adding the Cube Construction Test to the other two contributes little or nothing to the pooled reliability.

Table 1

Reliability Coefficients[1] Computed by the Split-half Method for: I. Grove Modification of the Kent-Shakow Formboard Series; II. Wechsler Block Designs Test; III. Cube Construction Test and Combinations of These Three Tests

School Grade	Tests								
	I				II	III	I+II	II+III	I+II+III
	N	T[2]	M[3]	C[4]	T	T	T	T	T
7	54	.82	.77[a]	—	.86	.58[b]	.88	.85	.90
8	56	.86	.85	—	.82	.65[b]	.92	.89	.92
9	62	.87	.77[a]	—	.78	.73[a]	.88	.88	.90
10	54	.88	.84	—	.87	.72[a]	.91	.91	.92
11	63	.89	.80	.86	.88	.58[c]	.90	.84	.90
12	63	.81	.79	—	.92	.64[b]	.90	.88	.92
9–12 (pool)	242	.84	.82	—	.86	.67	.90	.87	.91
7–8 (pool)	110	.84	.82	—	—	—	—	—	—

[1] Raised by the Spearman-Brown formula. Except where otherwise indicated the P.E.s for these correlation coefficients fall between ±.01 and ±.03. The exceptions have been designated by the following letters beside the coefficients: a = ±.04; b = ±.05; c = ±.06.

[2] T = time rate scores.

[3] M = move rate scores.

[4] C = time rate scores plus move rate scores.

Returning to the reliability of the formboard and the method used in its estimation, certain comments seem to be in order. The writer recognizes that the method used in this study may itself have lacked somewhat in complete reliability. To the extent that this was true, of course, the reliability coefficients obtained may have been lowered somewhat. The reliability of the formboard is, therefore, at least as high as these figures indicate and perhaps slightly higher. A more technically perfect method of studying the reliability of this test would have involved recording the performance of each subject by a movie camera calibrated for time. The expense of such an undertaking would be prohibitive, however, and the writer feels that the results here obtained justify the method used.

Conclusions

The following conclusions seem to be justified:

(1) The split-half method of determining reliability can be appropriately applied to performance tests of the type used in this study.

(2) The reliability of the Grove Modification of the Kent-Shakow Industrial Formboard Series is equal to or better than that reported for other manual and mechanical ability tests of comparable length and complexity. (See Bennett, G. K. and Cruikshank, R. M. *A summary of manual and mechanical ability tests (preliminary form)*. New York: Psychological Corporation, 1942.)

(3) Research on the validity of the Grove Modification of the Kent-Shakow Formboard Series would seem to be justified and the test is probably worthy of more adequate standardization.

(4) The use of move scores for this test does not seem to be justified in terms of reliability findings, especially when we consider the relative difficulty of counting moves.

Received May 9, 1946.

Influence of Leading Upon Readability of Newspaper Type *

Donald G. Paterson and Miles A. Tinker

University of Minnesota

Our survey of newspaper practice[1] revealed that the use of some leading is rather common practice. For body type the amount of leading employed ranged from ¼ point to 2½ points with ½, 1 and 2 points most popular. Fourteen of the 89 newspapers used no leading.

Although other factors may be involved to some extent, the main purpose of using leading is to improve readability of the textual material. In numerous studies conducted by the writers[2], readability has been measured in terms of speed of reading. In most of the studies reported, we also obtained readers' opinions concerning legibility of the printed text. The purpose in the present study is to determine the influence of leading on the readability of newsprint and to secure reader opinions of legibility for the material used.

A speed of reading technique was employed to measure the influence of various amounts of leading. The reading material consisted of Forms A and B of the Chapman-Cook Speed of Reading Test. In all comparisons Form A was printed in 7 point Ionic No. 5 newspaper type face in a 12½ pica line width set solid on newsprint paper stock as a standard. Form B, which was read after Form A, was typographically identical with Form A except for leading which varied in the different comparisons as follows: set solid, 1, 2, 3, 4, 5, 7, and 9 points.

Performance on Form B is equivalent to that on Form A on the average. However, since variation from equivalence sometines occurs[3], it is necessary to use a control group as a check.

Eight groups of 85 university students each were tested. In Group I, the control group, the typography was identical in Forms A and B, Thus a correction can be made in Groups II through VIII (the experimental groups) for whatever deviation occurs between Forms A and B of the control group. In Group II, and each of the successive groups,

* The writers are grateful to the University of Minnesota Graduate School for research grant to finance this study.

[1] M. A. Tinker and D. G. Paterson. War time changes in newspaper printing practice. *Journalism Quart.*, 1944, 21, 7–11.

[2] D. G. Paterson and M. A. Tinker. *How to make type readable.* New York: Harper and Brothers, 1940 (obtainable from the authors).

[3] D. G. Paterson and M. A. Tinker. *Op. cit.*, 1940, 132–145.

the material set solid was compared with material set with one of the amounts of leading specified above. By computing the differences between speed of reading text with each amount of leading and the text set solid, one can discover the influence of leading on relative readability of the typographical variations.

Table 1

Influence of Leading with Newspaper Type

Note: All reading material was set in 7 point No. 5 Ionic newspaper type in a 12½ pica line width on newspaper printing stock. Reading speeds for 0, 1, 2, 3, 4, 5, 7 and 9 point leading are compared with reading speed for material set solid as a standard. Plus (+) differences indicate faster reading than the standard. Number of readers = 85 university students in each test group, 680 in all.

Test Group	Test Form and Leading	Mean	PE_M	Difference Between Means in: Paragraphs*	Per Cent	r	$\frac{D}{PE_{diff.}}$**
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I	A, Set solid	15.44	.33	0.00	0.00	.90	0.00
	B. Set solid	15.74	.27				
II	A, Set solid	14.88	.29	+0.52	3.52	.77	2.83
	B, 1 pt. leading	15.70	.25				
III	A, Set solid	15.21	.29	+0.18	1.20	.89	1.35
	B, 2 pt. leading	15.69	.29				
IV	A, Set solid	14.82	.33	+0.41	2.74	.84	2.30
	B. 3 pt. leading	15.53	.28				
V	A, Set solid	15.26	.31	+0.73	4.82	.85	4.36
	B. 4 pt. leading	16.29	.29				
VI	A, Set solid	14.62	.31	+0.63	4.31	.80	3.37
	B. 5 pt. leading	15.55	.27				
VII	A, Set solid	15.12	.29	+0.42	2.76	.88	2.87
	B. 7 pt. leading	15.83	.29				
VIII	A, Set solid	15.52	.33	+0.32	2.08	.83	1.76
	B, 9 pt. leading	16.14	.30				

* The differences in column 5 are "corrected" by the amount of the difference between the mean scores of Form A and Form B of Test Group I which serves as a control group. The "correction" amounts to 0.30 paragraph for each test group comparison.

** $PE_{diff.}$ values in column 8 were computed on the basis of the standard formula for correlated measures.

The detailed statistical results are given in Table 1. It will be noted that any amount of leading from 1 to 9 points improves readability some-

what. Four points of leading, however, produces a highly significant improvement in readability; neverthless, in comparison with set solid the differences for 1 point leading are significant at the 3 per cent level, 3 point at the 6 per cent, 5 point at the 1 per cent, and 7 point at the 3 per cent. It appears, therefore, that 4 to 5 points of leading are optimum for readability of 7 point Ionic newspaper type face in a 12½ pica line width. There is a 4 to 5 per cent increase in reading speed. It is noteworthy that the largest amount of leading used (9 point) yields only slight improvement in readability. Apparently 1 point (3.5 per cent increase) leading is as good as any amount of leading except 4 and 5 point. The latter are more than is practical in most newspaper printing. From the standpoint of readability and what is practical, therefore, 1 point leading seems indicated for body type in newspaper printing when 7 point type is employed.

The reactions of readers to variations in leading were also obtained. An additional 221 college students ranked samples of the print according to apparent legibility (readability). In this part of the experiment, samples of 150 words for each amount of leading were mounted on cardboard and presented to the readers in a controlled manner.

The results derived from reader opinions of relative legibility are given in Table 2. The order of judgments is 4 and 5 point leading tied

Table 2

Newsprint Ranked According to 221 Reader Opinions of Relative Legibility for Variation in Leading with 7 Point: 12½ Pica Line Width

Leading	Mean Rank	S.D.	Rank Order
Set solid	7.31	1.22	8
1 point	5.95	1.73	7
2 points	4.57	1.80	5
3 points	3.30	1.65	3
4 points	2.96	1.42	1.5
5 points	2.97	1.59	1.5
7 points	4.01	1.96	4
9 points	4.87	2.46	6

for first place and then in order 3 points, 7 points, 2 points, 9 points, 1 point and set solid. Two things stand out in this table: (1) The readers consider most legible the text that was read fastest, i.e., text set with 4 and 5 point leading. (2) Readers consider text with any amount of leading to be more legible than material set solid. As pointed out by Tinker and Paterson[4],reader judgments of pleasingness tend to agree

[4] M. A. Tinker and D. G. Paterson. Reader preference and typography. *J. appl. Psychol.*, 1942, 26, 38–40.

with judged legibility. This implies that readers prefer newspaper body type with some leading rather than material set solid.

Summary and Conclusions

1. The purpose of this investigation is to determine the influence of leading on the readability of newsprint.

2. Speed of reading 7 point Ionic No. 5 newsprint in a 12½ pica line width set solid was compared with speed of reading text typographically identical except for leading which varied in the different comparisons as follows: set solid (control group), 1, 2, 3, 4, 5, 7, and 9 points.

3. All material with leading was read faster than text set solid.

4. In comparison with set solid material, the 4 and 5 point leading produced differences with the highest significance (beyond the 1 per cent level).

5. One point leading is as effective in improving readability as larger amounts of leading except for 4 and 5 points. Nine point leading is less effective than most of the lesser amounts of leading.

6. Text with 4 and 5 point leading is judged to be most legible. This agrees with the speed of reading results. Material set solid is judged to be least legible.

7. In view of practical considerations, readability and readers' judgments, 1 point leading is indicated for 7 point newspaper type in a 12½ pica line width. This would also hold for a 12 pica line width since unpublished data of the writers show that variations in line width from 12 to 12½ to 13 picas do not significantly affect readability of newspaper type.

Received April 27, 1946.

Prediction of Success in Learning Japanese *

S. B. Williams

Johns Hopkins University

and

H. J. Leavitt

Massachusetts Institute of Technology

Shortly before the end of the war, the authors undertook a study of methods of selecting students for the Marine Corps' Japanese Language School for enlisted personnel. Although the end of the war prevented the completion of the study, the preliminary data acquired seem sufficiently suggestive to be worth reporting. Informal inquiries have revealed no other comparable study.

The study was instigated by the excessive attrition rate in the school, approximating 80 per cent. Such a high rate is not particularly surprising in view of the rigorous curriculum, requiring acquisition of a speaking and reading mastery of the Japanese language in six months, and the lack of any standardized selection of students. Selection was made by interviewers who placed greatest emphasis on the applicant's General Classification Test (Army) score and specific motivation. College grades seldom could be used because most of the students were only high school graduates. Interviewers were school instructors, well aware of deficiencies in the method and anxious for improvement.

Because of time and circumstance, the tests chosen for experimental administration were exclusively of the paper and pencil variety. Plans did not materialize for constructing a phonetic test or a miniature-language test based on Japanese syntax. Most of the tests used are standard publications. A total of eight tests was administered which yielded 21 separate scores. The tests are: the U. S. Navy Officer Qualification Test, NavPers 16561, which includes Verbal Opposites, Mechanical Comprehension, and Arithmetical Reasoning; the U. S. Army Language Aptitude Test, W.P.Q.-1, X-1, a miniature-language test; the American Council on Education Psychological Examination for college freshmen, 1943

* The study was conducted at the Medical Field Research Laboratory, Camp Lejeune, N. C., with the cooperation of the staff of the Japanese Language School and the Marine Training Command. Collection and analysis of the data were done largely by N. Adler, C. L. Blair, N. J. Lipkin, and G. V. Mendola. Opinions expressed are those of the authors and are not to be construed as official or as representing the views of the Navy Department or of the Naval Service at Large.

edition; Anderson's Adult Placement Test (parts 1, 2, and 3 only); the Shipley-Hartford Retreat Scale; the NDRC Personal Inventory Form 4; a specially devised Symbol Digit Test in which the symbols were Japanese-like nonsense characters; and a specially devised Figure Recognition Test of visual memory for geometric forms. In addition, the AGCT score and the Army Mechanical Aptitude Test score were available from classification cards. A high level Verbal Analogies Test, recently constructed by the Bureau of Naval Personnel, was administered too late to be included in the criterion comparisons.

Two groups of subjects were tested. The *pre-school* group comprised students taking the tests immediately prior to school entrance. These were tested at the start of each month for each new class. After the third monthly group, subjects' school performance may not have been entirely free from the psychological effects of the war's end. Certainly some members of the later classes evidenced a marked shift in motivation. For this reason the present data are restricted to those students whose school outcome was decided prior to August 1945. This restriction considerably reduces group size but assures relative homogeneity among remaining subjects.

The *in-school* group included some of the students already enrolled in the school at the time the study was initiated. The particular students who were tested had already received two to five months of schooling, that is, they were in their third to sixth month. Inasmuch as most failures were detected by the beginning of the third month, the in-school group consists almost entirely of successes, although this particular group had fewer failures than average.

The pre-school group is an unbiased sample of entering students. The in-school group is a somewhat selected sample of advanced students. In analysing the results, the in-school group will be treated separately in some instances but for some purposes the groups will be pooled. It should be recognized that pooling is legitimate only on the assumption that specific tuition in the in-school group had no effect on test score. This seems a likely assumption, though it remains unproved.

After all eliminations there remain 56 subjects in the pre-school group and 19 in the in-school group.

The criterion employed is a pass-fail grade for the curriculum as a whole.

Results

Of the eight tests, two failed completely to differentiate passes from failures. These were the Personal Inventory and the Figure Recognition Test. The mean total test scores for the other six tests and for the AGCT are presented in Tables 1 and 2. Critical ratios are included

Table 1

Statistical Constants and Critical Ratios of Successes and Failures in Japanese Language School Tests Given at Beginning of Course

Tests	Successes			Failures			C.R.
	N	M	SD	N	M	SD	
Symbol Digit	9	73.8	19.6	47	60.1	17.8	1.9
Army Lang. Apt.	9	45.9	6.9	47	29.1	12.1	5.5
Adult Placement	9	54.0	9.7	46	42.9	12.9	2.8
Shipley-Hartford, Vocab. and Abst.	9	69.0	2.8	46	64.1	5.8	3.7
Navy Off. Qual.	9	63.2	13.8	46	51.7	22.4	2.0
ACE Psychol. Exam.	9	147.8	5.3	47	121.6	18.4	7.9
Army GCT	7	132.9	3.6	36	123.5	22.3	2.3

Table 2

Statistical Constants and Critical Ratios of Successes and Failures in Japanese Language School on Tests Given at Beginning of Course and in Third to Sixth Month after Course Began

Tests	Successes			Failures			C.R.
	N	M	SD	N	M	SD	
Symbol Digit	27	69.5	15.1	48	60.2	17.6	2.4
Army Lang. Apt.	27	43.6	9.2	48	29.1	11.9	5.8
Adult Placement	26	50.0	12.6	47	42.5	12.9	2.4
Shipley-Hartford Vocab. and Abstract	26	69.1	4.7	47	63.9	5.9	4.1
Shipley-Hartford Vocab.	26	34.6	3.4	47	31.9	2.7	3.3
Shipley-Hartford Abstract	26	34.2	4.2	47	32.0	4.6	2.0
Navy Off. Qual.	26	62.8	12.9	47	51.5	22.3	2.7
Navy Off. Qual. Arith.	26	13.2	3.1	47	12.1	2.7	1.6
Navy Off. Qual. Mechanical	26	18.6	4.5	47	16.5	3.9	1.9
Navy Off. Qual. Opposites	26	31.0	9.6	47	22.6	7.2	3.9
ACE Psychol. Exam.	27	144.6	16.0	48	121.3	18.2	5.6
ACE Lang. Score	27	92.4	11.3	48	75.1	11.3	6.3
ACE Quant. Score	27	52.2	9.7	48	46.0	9.6	2.7
Army GCT	25	131.8	8.7	37	123.6	22.0	2.0

to allow a rough comparison among tests. Though all are not presented, the score distributions give the same general picture.

The most discriminating tests are the Army Language Aptitude Test and the ACE. They are intercorrelated .63.

The former presents a novel language in miniature, with syntax similar to Romance Languages. It distributes the successful and unsuccessful students as shown in Table 3.

Table 3

Distribution of Army Language Aptitude Scores (Groups Pooled)

Score	0–9	10–19	20–29	30–39	40–49	50–59	Total
N Pass	0	0	1	5	9	12	27
N Fail	1	8	14	15	7	3	48

The latter is the usual scholastic test; it has two sub-sections scores: L (language) and Q (quantitative). To determine which of the two latter scores is the more discriminating of Japanese Language School grade, a breakdown was undertaken (Tables 2, 4, 5). The L score is definitely more effective than the Q score.

Table 4

Distribution of ACE Language Scores (Groups Pooled)

Score	40–49	50–59	60–69	70–79	80–89	90–99	100–109	110–119	Total
N Pass	0	0	1	1	4	14	4	2	26
N Fail	1	1	12	17	13	3	1	0	48

On the hypothesis that verbal tests are relatively more associated with the criterion, tests capable of similar breakdown were examined. In support of the hypothesis, the vocabulary score of the Shipley-Hartford Retreat scale is better than the abstraction score, and the Verbal Opposites section of the OQT is better than the Mechanical Comprehension and Arithmetical Reasoning sections. Therefore it seems highly probable that tests of "verbal intelligence" would be the best selectors of students for a school of this kind. The ACE Psychological Examination total score gives most evidence of satisfactory screening ability. Table 5 shows that a cut-off score (140) could be selected so as to screen 85 per cent of the failures with a false positive rate of only 11 per cent.

Table 5

Distribution of ACE Scores (Groups Pooled)

Score	70–79	80–89	90–99	100–109	110–119	120–129	130–139	140–149	150–159	160–169	170–179	Total
N Pass	0	1	0	0	1	0	1	12	9	2	1	27
N Fail	1	1	0	8	11	7	13	5	1	1	0	48

The mean ACE score of all students tested indicates a comparatively high level of scholastic aptitude. The mean of the failure group falls

at about the 70th percentile of the total male college freshman population. The mean of the success group is at the 91st percentile. For successful selection, then, the verbal tests must be oriented to about the highest fifth of the college population. The relative inefficacy of the AGCT probably can be attributed to its grossness in differentiating within narrow sub-groups of the total military population. It is unlikely, in fact, that any of our tests were of an optimally high level of difficulty.

It is not to be assumed that the criterion of pass-fail, used herein, is a pure measure of ability in the Japanese language. Unquestionably attenuating factors are present. Curricular time limits, for example, were far more severe than those which might obtain in a civilian institution in peace time and probably more severe than those in the Navy's school for language officers. However, the authors have been unable to ascertain the availability of comparable data from other schools.

Received April 27, 1946.

Effect of a Guidance Program on Emotional Development

G. M. Worbois

Detroit Edison Company, Detroit, Michigan

Many measurable as well as "intangible" results may accrue from a guidance program. It is, however, often difficult to demonstrate them to the critical observer. Measures used to show results due to guidance programs are frequently ambiguous, unreliable, and not objective. In the present study, an attempt has been made to overcome some of these difficulties. The purpose of the investigation was to secure an objective measure of one important psychological factor supposedly influenced by guidance, and to find its relationship to guidance procedures.

An intensive guidance program was carried on by the Flint Guidance Center (Flint, Michigan) with the support of the Rackham Sociological Research Unit. The project was initiated at the beginning of the ninth grade in a public intermediate high school. The entire grade was divided into two groups, comparable in sex, age, Kuhlmann-Anderson IQ, and Stanford Achievement test scores. One was designated as the "Experimental" group and the other as the "Control" group, there being 233 pupils in each of the groups.[1]

These groups were followed for three years and all students who remained in the public schools of the city during this time were retained in the project. The majority of students remained in the intermediate high school for one year and then transferred to a city senior high school. The comparisons reported in this study were made at the end of the three-year period.

The Experimental group was given intensive individual guidance. The emphasis was on educational-, vocational-, personal-, social-, and home-guidance, but whatever problem was significant for the adjustment of the student was given careful attention and appropriate treatment. Two well-trained and qualified counselors maintained unusually close contact with each of the students of this group. Psychological tests, interviews, occupational try-out experiences, home visitations, and what-

[1] The pupils were arranged alphabetically within thirteen groups, twelve successive ability groups as determined by Stanford Achievement test scores, and one group whose age was distinctly greater than that for the other pupils. Six "home-room" groups were drawn by taking the next name in order from each of the successive groups until one sixth of the names had been drawn. Three of these were arbitrarily designated as "Experimental" and three as "Control."

ever the ingenuity of the counselors could devise, were drawn on freely for the guidance of the students. The Control group received only the regular guidance program of the public high school, rather typical of most well-equipped city schools.

The particular area which it was desired to study in this investigation was what might be termed, "emotional adjustment." For this prupose the Luria technique (6) appeared to be promising. The Luria technique has been developed mainly during the last decade, although it employs psychological indices commonly used in other measures. The word-association method is combined with motor response in both hands, one of which is voluntarily directed. The other, the non-preferred hand, is not voluntarily directed, although non-voluntary movements are frequently found and are thought to be of significance.

The theoretical position underlying this technique may be best stated succinctly in Luria's own words (6): "We have shown that the marked disturbances in the motor system occur every time that the movements studied by us fall directly into the sphere of affect; we have come to the supposition that these disruptions arise precisely because the cortical apparatus is not in a condition to dominate the masses of excitation in the affect, and the movement unavoidably takes on an excitatory and destructive character" (p. 177).

Originally Luria seems to have been primarily concerned with the disorganization of the associative processes as the result of *strong* emotion. The method developed by Luria has also been used to detect *mild* affective reactions. As an example, the Luria method was used by Runkel to indicate the emotional disturbance accompanying deception (10). In another study, (8) the technique served to differentiate a group of young people, 17 to 25 years of age, some of whom were known to have conflicts in such things as "home," "companions of both sexes," "school or work," and "general outlook on life." It has also been found to be sensitive to *changes* in subjects accompanying an experimental program within an interval of less than a year. In one of these studies, (7) an experimental group took part in a series of discussions on personal and social problems. In another study, (11) the teachers were given detailed information concerning the children, their home, their particular problems, etc., with suggestions concerning best methods of treatment. In both studies significant changes were found in the Luria scores, as shown by increasing differences between experimental and control groups.

The Problem

The question proposed in the present study may be stated: After three years of guidance, is there a difference in Luria scores between the

Experimental and Control groups? It is assumed in advancing this problem that conflicts over personal, educational, vocational, or social problems can be solved, or at least mitigated by counseling.

Uncertainty of the student as to whether he was in the right curriculum for his abilities and interests might lead to a conflict. Testing his aptitudes and advising him concerning their implications might in many cases be expected to lead to a satisfactory solution. In turn this might be expected to reduce his affective reaction in these areas, and would possibly be detected in lower Luria scores.

Another type of problem is that in which the subject is limited to two alternatives, neither of which seems to be desirable. The counselor, understanding the background of the problem and the reason for the anxiety of the subject, may assist the student to make a satisfactory selection. In many such cases it might be expected that there would be an affective tone, especially when some action must be taken. If the effect of the counseling has been to provide a satisfactory solution it seems reasonable that the unpleasant affective reaction would be reduced.

Another type of anxiety is that concerning personal problems. Here the source of the disturbance seems to be largely subjective. The subject becomes "worried" over personality or social traits which in some way have been interpreted as "bad." The result of counseling in such cases might be thought to lead to a more objective evaluation by the subject of his own traits. Reduction in the subject's disturbance over these traits might be expected, and this in turn might be indicated by lowered Luria scores.

In several ways, then, it might be expected that counseling would effect a more objective and analytical approach to problems of the kind indicated. This is true not only in the specific problems on which the subject has had advice, but also in the manner of his approach to problems, in breaking them down into their essential parts, and in formulating a systematic evaluation of possible solutions. For example, one of the teachers is reported to have said that the Experimental students were more free to discuss their problems than were the Control students. One interpretation of this kind of statement is that these students have taken a more objective attitude toward their problems, and are willing to accept advice or assistance with them. Even though such an observation might not be systematically made it seems to be in agreement with the expected results of counseling.

Not all counseling would be expected to resolve conflicts, of course. Some problems might be raised, the importance of which the subject had not realized. It might be expected, however, that when such problems are raised the subject is given the information and guidance necessary to

make an analytical evaluation, and a more objective decision. It also might be expected that the conflict would be less than if the subject had come upon the problem without information and guidance.

It might reasonably be expected, therefore, that in general, guidance would encourage a more objective, analytical, and impersonal approach to educational, vocational, personal, and social problems. If this assumption is valid it might be expected also that subjects receiving guidance of this kind would show less conflict over the problems in these areas as measured by the Luria technique.

Definitions

To make clear the type of psychological processes proposed for measurement it might be well to set down some working definitions. *High conflict* has been defined by Morgan and Ojemann (8) as: "Those individuals who have personal and social problems which they have not solved, which they feel incapable of solving, and toward which they have developed an attitude of dislike to the extent that they tend to repress or inhibit a discussion of them."

The same authors describe individuals having *low conflict* as: "Subjects in this group give evidence of adopting a more objective attitude toward their problems. They give evidence of feeling relatively secure to the extent that they are willing to discuss their problems. They give few indications that they fear new data but rather seek them and attempt to integrate them into their attitudes."

These definitions are not intended to be definitive, or static. They are merely set down as working directions to distinguish the psychological processes being studied.

Apparatus and Procedures

A polygraph which recorded a fine line by means of an electrical contact with carbon-backed paper was arranged with four pens. Closed pneumatic systems activated the pens, the two for recording motor responses adjusted to register very sensitive movements. One was connected with a tambour on which the subject rested his preferred hand; another was connected with a similar tambour for the non-preferred hand. One pen was used as a time line and indicated one-second intervals by the breaking of an electric circuit. The fourth pen was used to indicate the interval between the word spoken by the experimenter and the subject's verbal response. This was activated by a tambour controlled manually by the experimenter.

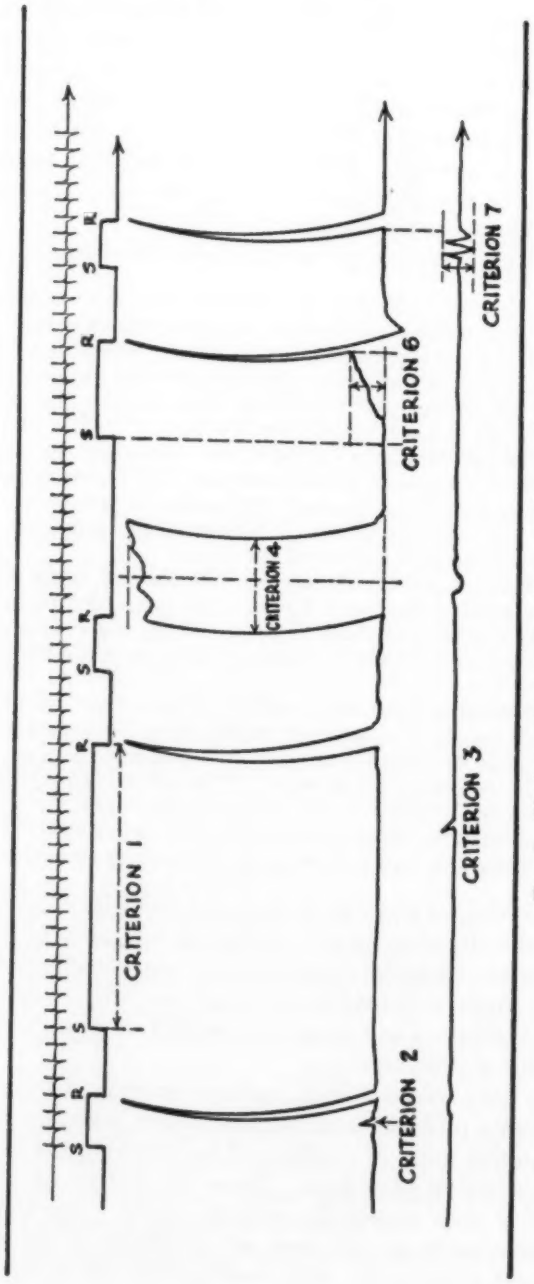

FIG. 1. Various criteria secured from Luria records.

The tambours on which the subject rested his hands were mounted on $\frac{5}{8}$" panel, sufficiently high from the table so that there was no support beyond the wrist, and so that the entire length of the forearm as far as the wrist could be rested on the table. A movable screen was placed in front of the entire apparatus with the exception of the panel on which the tambours were fastened. After the subject had been seated and had placed his hands on the tambours the screen was moved forward so that the subject could not see his hands.

After the subject was seated the following instructions were given:

We are arranging a new kind of testing apparatus and would like to have your help in trying it out. Since there is not enough time to take all of those in your class a few have been selected without any reference to their school work, personal abilities, or problems. You are one who has been selected and we would like your co-operation in helping work out this new device. It is rather interesting and you will probably enjoy it.

You will notice two tambours on the table. Place the first two fingers of each hand lightly on the discs. (Demonstrate) I shall read a word to which you are to respond with the first word that comes to your mind, and at the same time press with your right hand (if right-handed) on the disc. With the other hand keep a constant pressure all the time.

Remember, reply as quickly as you can with the first thing that comes to your mind after hearing the word I read. No particular reply is right or wrong; merely give what you think of first. At the same time press on the right disc. Also remember to keep pressing lightly with the left hand. (Reverse if left-handed)

Keep alert all the time because you will not know what word will be coming next. There will be about a hundred words, which will take about fifteen minutes. There will be some sound from the machine (demonstrate) which turns the paper for recording the answers. This will not affect these tambours which you are touching, however, and may be disregarded.

Since some people are confused in thinking of words quickly if they watch their hands, we place this screen in front of you and over your hands.

None of the subjects had had previous contact with the experimenter. He assumed a friendly attitude but confined all comments as much as possible to the printed instructions so that conditions would be as constant as possible. The words were read in a natural tone with a minimum of emphasis and inflection. The testing was carried on in a room relatively free from outside distractions.

The words were selected from various sources. Personnel records were read to locate problems and counselors were asked for words which designated common areas of conflict. The following list of words was devised for the testing procedure. Those marked by an asterisk were added because of their supposedly neutral nature. The first ten words were used as practice items and were not considered in any of the measures. They were given in the order reading from left to right.

tree*	friend	homely	city*	father
blues	soldier	death	guilty	insane
petting	sister	cheat	sex	pimples
dreams	loafing	dirty-story	suspicion	faint
snubbed	pencil*	glasses	queer	sneak
work*	mathematics	examination	confess	fear
worry	home-life	lie	money	unhappy
private	sin	pay	dating	inferior
blush	love	secret	white*	showers
study-period	factory	brother	unwanted	trouble
teacher	smoke	grades	crying	truant
failure	step-mother	timid	sarcastic	crazy
welfare	window*	debts	drunk	health
dirty	stern	manners	mother	jealous
nagging	divorce	negro	desk *	sissy
affectionate	marriage	overweight	old-fashioned	self-conscious
tobacco	step-father	gamble	assembly-line	machine-operator
paper*	lipstick	touchy	complexion	counselor
summer*	study	skill	office-work	tardy
janitor	gymnasium	anxiety	straight*	allowance

Little attention has been given to re-test reliability and validity of the Luria technique in this study since considerable research is available elsewhere (2), (3), (8), (9). The present study is not an investigation of the technique itself, but is primarily an application of the technique.

On the other hand, objectivity, accuracy of measurement, and rater reliability were considered essential for the study. Criteria were selected which required a minimum of subjective evaluation.

Of the five criteria used in the study (see Table 2) there was only one which required the rater to judge whether the response was significant or not. This was criterion number 2, definite node in the preferred hand. This was defined carefully so that there would be little disagreement between raters as to which responses showed a conflict.

To test this agreement two raters independently and with no special training marked the responses considered to be significant on the eight records reported below in the "Preliminary Study." One rater marked 133 words as showing a conflict according to this criterion, and of these 127 were also marked by the second rater. This gives 95 per cent agreement. Since one rater marked more words than the other, the total number of words marked was divided into the total number of agreements, giving 90 percent agreement. There seems to be sufficient reliability, accordingly, to warrant the use of this criterion.

The "reaction time," or time between the stimulus word and the sub-

ject's response, was easily computed from the time line on the polygraph paper to within ± 0.5 sec. Criterion 4, Table 2 (prolonged motor response) was determined by means of bow spring dividers. This means of comparing distances involved a relatively small error. The widths of the two preceding and two following responses at the mid-points (defined by half the distance from lowest to highest points of the curve) were added together and compared with half the width of the designated response.

Criterion 6, Table 2 (mean deviation in millimeters in curve for preferred hand before motor response) was determined as demonstrated in Figure 1 by means of a transparent tape calibrated in millimeters. This was found to be a more accurate method than the Vernier calipers for the purpose, and involved an error of measurement less than ± 0.5 mm.

The Subjects

Twenty-four subjects were selected from the Experimental, and twenty-four from the Control groups, twelve girls and twelve boys from each. There were thus 48 subjects in the study. The names of all the boys and the girls remaining in the two groups in the fall of 1942 were arranged in alphabetical order. Form each of these four groups a random sample of twelve subjects was selected by means of a table of random numbers (5).

The samples were drawn to be random samples with reference to Luria scores. Since they are random samples, however, other factors should also be randomly distributed. As can be seen in Table 1, the mean Kuhl-

Table 1

Comparison of Kuhlmann-Anderson IQ for Random Samples with Total Experimental and Control Groups

	Total Group			Random Sample			
	N	M	σ	N	M	σ	Diff.
Experimental							
Boys	64	101.4	12.7	12	102.4	9.4	+1.0
Girls	78	100.6	10.9	12	99.5	9.0	−1.1
Total	142	101.0	11.8	24	101.0	9.3	0.0
Control							
Boys	57	101.0	12.5	12	102.3	9.1	+1.3
Girls	70	103.9	12.3	12	104.7	9.8	+0.8
Total	127	102.6	12.5	24	103.5	9.5	+0.9

mann-Anderson IQ'S of each of the four samples approaches very nearly the average IQ of the group from which it was drawn. The mean IQ for the total Experimental group was 101.1, the same as the mean IQ for the

sample drawn from this group. The mean IQ for the total Control group was 102.6, while that for the sample was 103.5

Preliminary Study

Two boys and two girls were selected by one of the counselors on the staff of the Flint Guidance Center from the Experimental group because they were thought to have conflicts of the kind indicated in the "working definition" above. Two boys and two girls were also selected because they were thought to be "non-conflict" students according to the definition given. The Luria test was given to these eight cases to find how well it distinguished between the students the counselor recognized as "conflict" and those thought to be "non-conflict." It was hoped also to find some indication of what measures might profitably be secured from the Luria records.

Seven measures were secured from these eight records. These are described in Table 2. In addition to these criteria, reaction time on the total list of 90 words and the reaction time on the words not showing a conflict as judged by the composite criterion, number 5, Table 2, were recorded. It appeared that little sex difference was present in any of these measures so that a comparison of the four "conflict" and four "non-conflict" cases would be sufficient. As can be seen in Table 2 the criterion which appeared to differentiate the two groups best was number 2, definite node in the motor response of the preferred hand.

With the exception of criteria 1 and 4 the differences are in favor of less disturbance for the "non-conflict" students. The fact that more of the differences are not statistically significant must be interpreted in terms of the small number of cases, and the purpose of this preliminary study. Criteria numbers 3 and 7 did not appear especially promising because of the small number of incidences. It was decided to use criteria numbers 1, 2, 4, 5, and 6 for the study proper.

Analysis of Results

The mean reaction time for the Experimental sample of 24 cases was $3.22 \pm .13$ sec. on all words for which there was no indication of conflict according to the composite criterion. For the Control sample it was $3.31 \pm .12$ sec. The mean reaction time for the Control group on all 90 words (including those on which there was indication of conflict) was 1.00 sec. longer than the reaction time on only words showing "no conflict." For the Experimental group the reaction time for the entire list of words was 0.85 sec. greater. These reaction times are similar to those in other studies using a manually controlled measure of reaction time, although it is greater than the reaction time reported in those studies using a voice key and chronoscope as a measure (**2**), (**3**), (**10**).

Table 2

Comparison of Four "Conflict" and Four "Non-Conflict" Cases on Seven Measures Secured from the Luria Records

Criteria	Conflict M	Non-Conflict M	Diff.	t_0*
1. Number of words having reaction time of 15 sec. or more.	2.5	5.8	+ 3.3	0.48
2. Words showing conflict as judged by definite node in line for preferred hand (upright) between giving stimulus word and response. There must be a marked acceleration, point of inflection, and marked deceleration between stimulus and upward sweep of motor response. Size of node is not especially significant.	24.0	9.3	−14.7	4.38
3.** Words showing conflict as judged by definite node (upright or inverted) in line for non-preferred hand between any stimulus word and the following word except where it coincides with the movement of the preferred hand.	8.3	2.5	− 5.8	1.92
4. Words for which the width of the stroke at the midpoint (defined by lowest and highest points) is more than twice the average of the two preceding and the two following strokes.	1.5	2.0	+ 0.5	0.25
5. Words showing conflict as judged by any one or more of the criteria 1, 2, and 4 above.	25.5	13.8	−11.7	1.01
6. Average deviation in mm. above the level at which the stimulus was given in the preferred hand, and preceding the motor response.	5.1	2.4	− 2.7	0.76
7.** Average deviation in mm. of the range in the non-preferred hand, either upward or downward from the point at which the stimulus was given to point at which the following word was given.	1.2	1.0	− 0.2	0.29

* t_0 is the obtained value of t. The theoretical value of t is 2.44.

** Not used in study proper.

The Control students had a mean conflict on 24.58 of the 90 words as judged by a definite node in the curve for the right hand. (This is illustrated in Figure 1.) The Experimental students had a mean conflict on only 14.17 of the 90 words, the difference being statistically significant at the one per cent level of confidence according to the t-test. (See Table 3.) The ratio of the difference to the standard error of the difference is also given for comparison.

The Control sample had a larger number of prolonged motor responses in the preferred hand as judged by criterion number 4, Table 3. The dif-

ference, 1.19, between the two groups is statistically significant beyond the five per cent level of confidence.

The mean deviation in mm. (criterion number 6, Table 3) in the curve for the preferred hand was 0.60 mm. greater for the Control sample

Table 3

Luria Scores for 24 Experimental and 24 Control Subjects

	Control		Experimental		Differences			t_t***	
Criteria Number*	M	σ	M	σ	Diff.	$\frac{\text{Diff.}}{\sigma_{\text{diff.}}}$	t_0**	(5%)	(2%)
1	3.00	4.41	3.00	4.23	0.00	0.00	0.00	1.96	2.32
2	24.58	14.25	14.17	12.63	−10.41	2.68	2.62	1.96	2.32
4	1.42	1.93	0.23	1.62	− 1.19	2.32	2.27	1.96	2.32
5	26.00	12.70	16.46	12.41	− 9.54	2.63	2.52	1.96	2.32
6	2.28	1.62	1.68	0.93	− 0.60	1.55	1.52	1.96	2.32

* For description of criteria see Table 2.

** t_0 refers to the obtained value of t.

*** t_t refers to the theoretical value of t required to be significant at the level of confidence indicated.

than for the Experimental sample. The t-value for the difference is 1.52, which is significant at about the eleven per cent level of confidence. It fails to be reliable statistically even though it is consistent with the other findings.

According to the composite criterion (number 5, Table 3) the Control sample had a conflict on more words (26.00) than did the Experimental sample (16.46). The difference is statistically significant at the two per cent level of confidence according to Fisher's t-test.

As might be expected from the preliminary study there was no difference between the two groups in number of words for which the reaction time was 15 sec. or more, the mean being 3.00 in both cases.

Discussion and Conclusions

1. The group of students subjected to an intensive guidance program showed significantly less "emotional conflict" than did the Control students. Before concluding that the differences were due to the guidance procedures it would be necessary to eliminate other possible explanations. Reasonably thorough attempt has been made to discover if there have been other factors to which differences in the Luria scores could be attributed. Some of these are discussed below.

One important point is whether there was any difference between the two groups before the experimental guidance program, i.e. at the begin-

ning of the ninth grade. There seems little reason to suspect initial differences because of the method by which the groups were selected. Names of all the pupils in the school at the end of the eighth grade were arranged in alphabetical order within ability groups (thirteen groups divided according to achievement test scores and age). Six groups were drawn simply by taking the next name in order from each of the successive ability groups until one sixth of the names had been drawn. There resulted six equivalent groups, three of which comprised the Experimental and three the Control group. No differences between the groups, except in those factors associated with the first letter of the surname, could be expected. It seems reasonable to assume, accordingly, that the Experimental and Control groups were originally similar on measures used in this study.

The next critical question is whether there has been selective elimination in either group. That is, have the students with more emotional conflicts dropped out of the Experimental group? The opposite trend, however, would be expected. It is probable that the guidance procedures would tend to encourage students with conflicts to remain in school. This is borne out by the fact that the mean IQ of this group is slightly below (*see* Table 1) that of the Control group, although both groups had the same mean IQ at the beginning of the project. Other evidence of a descriptive sort, too voluminous to be included, shows that the counselors frequently aided students to overcome their problems and thus remain in school.

The testing conditions do not appear to account for the differences. The experiment was carried on in a room fairly well isolated from distraction and was unfamiliar to both groups. The experimenter was unknown to all subjects and assumed an attitude intended to be constant throughout. This was facilitated by the fact that he made it a point of not infroming himself as to which group a subject belonged during the experimentation or while he was working with the records. Both groups were tested at approximately the same time over a period of about a week without making an attempt to test one group before the other. As far as the experimenter could determine there were no unusual events in the school or community program to influence one group more than the other.

2. The results of this study indicate that emotional adjustment to things represented by the stimulus words can be modified over a three-year period by an appropriate guidance program, and that these trends can be detected by reliable psychological indices. Just how these apparent changes are related to emotional adjustment of other kinds and according to other criteria has not been included in this study.

Received March 26, 1946.

References

1. Burtt, H. E. Motor concomitants of the association reaction. *J. exp. Psychol.*, 1936, **19**, 51–64.
2. Gardner, J. W. An experimental study of the Luria Technique for detecting mental conflict. *J. exp. Psychol.*, 1937, **20**, 495–505.
3. Houtchens, H. M. A study of mental conflict in delinquent and non-delinquent boys. *J. Juvenile Res.*, 1935, **19**, 180–192.
4. Huston, P. E., Shakow, D., and Erickson, M. H. A study of hypnotically induced complexes by means of the Luria Technique. *J. gen. Psychol.*, 1934, **11**, 65–97.
5. Lindquist, E. F. *Statistical analysis in educational research.* New York: Houghton Mifflin Co., 1940. Pp. 266.
6. Luria, A. E. *The nature of human conflicts.* Trans. by W. H. Gantt. New York: Liverwright, 1932. Pp. 431.
7. Morgan, Mildred I., and Ojemann, R. H. The effect of a learning program designed to assist youth in an understanding of behavior and its development. *Child. Develpm.*, 1942, **13**, 181–194.
8. Morgan, Mildred I., and Ojemann, R. H. A study of the Luria Method. *J. appl. Psychol.*, 1942, **26**, 168–179.
9. Olson, Dorothy M., and Jones, V. An objective measure of emotionally toned attitudes. *J. genet. Psychol.*, 1931, **39**, 174–196.
10. Runkel, J. E. Luria's motor method and word association in the study of deception. *J. gen. Psychol.*, 1936, **15**, 23–37.
11. Ojemann, R. H., and Wilkinson, Frances R. The effect on pupil growth of an increase in teacher's understanding of pupil behavior. *J. exp. Educ.*, 1940, **8**, 143–147.

Further Application of Simplified Scoring of the Bernreuter Personality Inventory

David C. McClelland

Wesleyan University

In a previous article (**4**) the writer reported a method of scoring the Bernreuter Personality Inventory which greatly economized on the labor involved in scoring it by reducing all weights of +3 or more to +1, all weights of −3 or more to −1 and disregarding weights from +2 to−2. These shortened scores correlated so highly (.95 or more) with B1N, B4D, F1C and F2S that it was suggested they be used in place of the full scores. In the same issue of the journal in which the article appeared another paper written by Kempfer (**3**) reported that he had tried the same method of simplifying the scoring (his Key II) but had obtained correlations with the full scores of only .69 to .91. This immediately raised the question as to what caused the marked discrepancy in results.

The first possibility was that some error had been made in scoring or calculations. The data used in the writer's report were rechecked and and found correct. Unfortunately, Kempfer's data were no longer available for checking. On the assumption that his too would have proved correct, this leaves the possibilities that the writer's data showed spuriously high correlations because of some selective factors in the sample of college men used, that the correlation of full and short scores is high for some groups (college men) and not for others (High School boys), or that the shortened scores are not as reliable as the full scores making possible wide variations in correlations obtained. Even had Kempfer's data not appeared it would seem desirable to test the representativeness of the correlation sample, the reliability of the short scores, and their correlation with full scores in other groups. These are the purposes of this paper.

Representativeness of the Sample

The simplest way to approach this problem is to compare the full score means and standard deviations of the correlation sample with the like means and standard deviations of Bernreuter's published norms. This is done in Table 1. For B2S, B4D and B1N (sample B) the means and SD's of the sample closely approximate those of the population. In the case of F1C and F2S the SD's are comparable but the means differ significantly. Still it is doubtful whether this would influence the correlations for two

Table 1

Comparison of Full Score Means and Standard Deviations of College Men in the Correlation Samples and in Bernreuter's Norms, Showing in Each Case the Correlation Between Full and Shortened Score in the Sample

	B1N			B2S		B4D		F1C		F2S	
	Bernreuter Norms	Correlation Sample A	Correlation Sample B	Bernreuter Norm	Correlation Sample	Bernreuter Norm	Correlation Sample	Bernreuter Norm	Correlation Sample	Bernreuter Norm	Correlation Sample
N	656	114	129	658	114	631	114	273	114	273	114
Mean	−57.3	−71.2	−49.7	27.0	19.9	45.9	40.4	−51.5	−30.7	−25.9	−45.4
SD	82.2	68.9	56.2	52.8	43.8	65.6	59.5	83.6	74.3	59.4	61.1
Correlation between full and shortened score		.966	.963		.844		.955		.975		.955

Table 2

Comparison of Full Score Means and Standard Deviations of College Women in the Correlation Samples and in Bernreuter's Norms, Showing in Each Case the Correlation Between Full and Shortened Score in the Sample

	B1N		B2S		B4D		F1C		F2S	
	Bernreuter Norm	Correlation Sample	Bernreuter Norm	Correlation Sample	Bernreuter Norm	Correlation Sample	Bernreuter Norm	Correlation Sample	Bernreuter Norm	Correlation Sample
N	544	100	391	100	396	100	144	100	144	100
Mean	−42.8	−39.8	6.8	−1.5	30.6	27.9	8.7	17.1	−31.1	−24.5
SD	76.8	73.1	54.2	57.8	61.8	62.9	75.1	92.3	56.4	53.7
Correlation between full and shortened score		.971		.942		.946		.968		.901

reasons, (1) because the two samples used for B1N differed widely both in mean and SD without affecting the correlation appreciably, and (2) because the SD is much more apt to affect the correlation than the mean. Only if the range were higher for the sample would one expect a spuriously high correlation.

Correlation of Full and Short Scores for College Women

To find out if the high relationship held only in the group of college men previously tested, an exactly comparable study was made on a group of Bryn Mawr college women.

Table 3

Percentile Equivalents in Bernreuter's Norms for Given Short Scores of College Women (N = 100)

Percentile Standing in Correlation Sample	B1N		B2S		B4D		F1C		F2S	
	sB1N	%ile equiv.	sB2S	%ile equiv.	sB4D	%ile equiv.	sF1C	%ile equiv.	sF2S	%ile equiv.
99	32	99	27	99	27	99	47	99	25	99
95	22	95	18	95	24	96	47	98	18	94
90	18	92	13	89	21	92	30	96	13	88
80	11	82	9	75	17	80	17	85	9	82
70	5	71	5	65	13	69	11	71	5	76
60	0	62	2	56	9	60	7	62	3	67
50	− 3	50	− 1	45	6	47	4	54	0	54
40	− 6	40	− 4	35	2	35	1	46	− 3	40
30	−11	29	− 6	23	− 2	25	− 3	35	− 6	32
20	−15	18	−11	9	− 5	17	−11	18	− 8	19
10	−21	13	−14	5	−10	10	−17	8	−12	13
5	−23	8	−17	3	−15	3	−23	3	−15	9
1	−31	2	−22	1	−23	1	−31	1	−24	3
M	− 2.2		− 0.4		5.2		4.7		0.3	
SD	14.3		10.7		12.0		16.9		10.3	

Note: The "s" preceding the scale designation (e.g. sB1N) means the simplified score for that Bernreuter scale.

Table 2 presents the results showing both the correlation between full and short score and the representativeness of the sample as compared with Bernreuter's published norms for college women (2). In this case the sample appears to approximate the population more closely and the correlations remain as high as they were for college men, the only difference being a reversal of the degree of relationship between F2S and B2S. With the men F2S is more closely correlated with the full score; with the women it is B2S. It is not possible to say from these data what is causing the reversal, but it would seem wise for the time being to get both scores

and use them as a check on each other since they measure closely allied traits.

As in the case of college men regression equations may be derived from the data presented so that the most probable full score may be calculated from a given short score. Another simpler procedure[1] is to construct tables which will give the percentile standing in Bernreuter's norms of any given short score. Tables 3 and 4 present these data for college women and college men respectively.

Table 4

Percentile Equivalents in Bernreuter's Norms for Given Short Scores of College Men (N = 114)

Percentile Standing in Correlation Sample	sB1N	%ile equiv.	sB2S	%ile equiv.	sB4D	%ile equiv.	sF1C	%ile equiv.	sF2S	%ile equiv.
99	24	94	27	96	32	99	32	97	30	98
95	18	88	20	87	28	93	21	92	14	88
90	14	85	18	83	24	85	17	90	10	81
80	6	70	13	72	20	76	10	84	6	69
70	2	66	11	65	17	67	3	80	3	60
60	− 3	56	8	54	12	55	− 1	70	− 1	50
50	− 7	48	5	51	9	46	− 5	62	− 5	40
40	−11	40	2	39	6	37	− 9	54	− 8	33
30	−16	29	0	27	1	29	−15	44	−11	27
20	−20	15	− 4	18	− 3	19	−21	30	−19	12
10	−27	7	− 7	9	− 9	11	−29	18	−27	3
5	−31	5	−10	5	−13	6	−33	12	−33	1
1	−35	2	−17	3	−21	2	−41	4	−39	1
M	− 6.9		5.1		8.4		− 5.7		− 6.1	
SD	14.3		9.3		12.5		16.6		14.1	
Means and standard deviations based on larger samples including the above 114 cases										
N	301		300		211		177		212	
M	−6.0		4.5		8.1		−7.3		−9.2	
SD	14.2		8.6		11.5		16.9		14.8	

Note: The "s" preceding the scale designation (e.g. sB1N) means the simplified score for that Bernreuter scale.

The method used in constructing Tables 3 and 4 was to find the full and short scores having the same percentile standing in the same sample (in this case the correlation sample) and to convert the full score to its equivalent percentile rating in Bernreuter's published norms. Since the

[1] The writer is indebted to Dr. Edward E. Cureton for suggesting this procedure.

correlations are high, it may reasonably be assumed that the same men will obtain full and short scores having equal percentile standing within the same sample. Since Tables 3 and 4 take only representative points throughout the sample, a great deal of interpolation will be necessary. Consequently, anyone planning to use this method of scoring should use the points given to construct a curve from which more accurate Bernreuter percentiles can be read off for any short score. At the bottom of Table 4 are listed the means and standard deviations for each scale not only in the correlation sample but also in the larger group of college men tested on whom the percentile norms previously reported (**4**) were based. These Tables should make a short score for a college man or a college woman immediately meaningful in terms of other research done with the Bernreuter Personality Inventory in which full scoring was used.

Reliability of Short Scores

Are the simplified scores less reliable than the full scores? Table 5 gives the answer by comparing the split-half odd-even reliabilities ob-

Table 5

Comparative Odd-even Reliabilities for Simplified and Full Scores (after Bernreuter (1))

	N	B1N	B2S	B4D	F1C	F2S
Simplified Scoring						
College Men	100	.87	.81	.83	.82	.84
College Women	75	.88	.82	.89	.82	.88
Full Scoring	70	.91	.92	.89	—	—
Full Scoring	100	—	—	—	.86	.78

tained with both sets of scores. It is clear that there is no loss in reliability resulting from using the shortened scores. In some instances there is even an apparent gain.

It has been demonstrated that the shortened method of scoring can be used with no loss in reliability and with close approximation of what Bernreuter was measuring both among college men and college women. There is doubtless some loss in validity from not using the full scale but since only the most heavily weighted items are used in the simplified scale, this loss cannot be serious. In a day when the overall value of questionnaire measures of personality is being increasingly doubted, it is important that the information such a test *can* usefully provide be available quickly and economically. Otherwise college advisers, faced with the expense of obtaining essential but not widely useful information, are likely to decide to do away with the tests. To draw an analogy from

medicine, doctors nearly always take a patient's temperature but if it were a lengthy or expensive process, they would omit it more often, not because they didn't think it important but because the information it gives would not be worth the expense. So it is with test scores.

The chief problem in the use of short scores is that the other method of scoring has been used so long that it is confusing to use a new one. Results cannot be exactly compared with previous results. To some extent this problem can be met by the tables of equivalent percentiles. To a greater extent it should be met by a growth of research using simplified scores since the latter are so much simpler, quicker and less expensive to apply. To avoid confusion it would seem preferable to label short scores with an "s" (i.e. sB1N, sB2S, sB4D, sF1C, sF2S) so that readers will know which method of scoring has been used.

Summary

This paper presents a further analysis and extension to new data of the simplified method of scoring the Bernreuter Personality Inventory previously reported (**4**), in which the answers weighted -3 or more by Bernreuter are counted and subtracted from the number of answers weighted +3 or more. As in the case of the college men on whom the original study was done, application of simplified scoring to college women yielded scores which correlated very highly (.94 or above for B1N, B2s, B4D, F1C and .90 for F2S) with the full scores. In both instances the correlation with either F2S or B2S was lower than the rest suggesting that both scores should be obtained to check against each other.

The usefulness of the shortened scoring was further attested by data showing that the samples used in obtaining the correlations were representative, that reliabilities for short scores are as high as for full scores, and that short scores may be readily converted into full score percentiles by means of tables provided. If, as seems likely because of their much greater economy, short scores are more widely used, it is suggested that the letter "s" precede each scale designation (e.g., sB1N) to avoid confusion with full score results already published.

Received April 1, 1946.

References

1. Bernreuter, R. G. *Manual for the personality inventory.* Stanford University: Stanford University Press, 1938.
2. Bernreuter, R. G. *The personality inventory: tentative percentile norms.* Stanford University: Stanford University Press, 1938.
3. Kempfer, H. Simplifying the scoring technique of the Bernreuter personality inventory. *J. appl. Psych.*, 1944, **28**, 412–413.
4. McClelland, D. C. Simplified scoring of the Bernreuter personality inventory. *J. appl. Psych.*, 1944, **28**, 414–419.

Braille Edition of the Minnesota Multiphasic Personality Inventory for Use with the Blind *

Orrin H. Cross

University of Alabama

In view of the common conception of the blind as socially maladjusted individuals (3,8), it appeared desirable to adapt a new test of personality for use with the blind. Accordingly, the Minnesota Multiphasic Personality Inventory (MMPI) was transcribed into braille in order to determine the feasibility of its use with the blind.

Review of the literature yielded but two studies which used standardized measures of personality on blind subjects, both by Brown (1,2). He found evidence that the blind females were more introverted and had a greater tendency to be neurotic than sighted females, while the blind males were less introverted but more neurotic than sighted males.

Procedure

The sample was drawn from the blind accessible in the vicinity of the University of Minnesota where the study was conducted. It included subjects from a cooperative dormitory, a sheltered workshop, a state school for the blind, and a number of subjects secured through the cooperation of the Minnesota State Department of Rehabilitation and the Sight Saving Departments of the public schools of two different cities (see Table 1).

The grouping of "16–25" vs. "26–up" was selected on the basis of the fact that the authors of the inventory divided their norm groups into 16–25, 26–35, etc., up to 56–65. Since the group tested in this study had insufficient numbers in the older groups to warrant separate treatment, they were lumped into one "older" group.

The grouping by "Time of Onset" was based on a personal letter from Dr. S. P. Hayes to the author ". . . in the schools, we . . . are accus-

* The author wishes to acknowledge the aid and advice of Prof. Donald G. Paterson in choosing the thesis problem and planning the study on which this article is based and to Dr. S. R. Hathaway who aided in securing permission to have the MMPI transcribed into braille and encouraged the author to make the study. Special credit should be given to Dr. Paul E. Meehl for his constant aid and patient help in executing the research. In addition Donald Dabelstein, Miss Gladys M. Fairbank, Melvin Fredell, J. Arthur Johnson, J. C. Lysen, C. Stanley Potter, and Byron Smith helped in securing subjects. Mrs. Mildred B. Cross helped score and record the tests. The author's M.S. thesis with the same title is on file in the University of Minnesota Library.

Table 1

Distribution of Blind Sample by Sex, Age, Time of Onset, Amount of Blindness, Education, and Employment Status

	Age at Time of Testing		Time of Onset		Vision		Education**		Employment Status	
	16–25	26–up	Before Age 4	After Age 4	Totally Blind	Some Vision*	Residential School	Public School	Employed	Unemployed
Males	12	13 (to 51)	11	14	16	9	10	6	11	6
Females	12	13 (to 66)	18	7	8	17	9	8	8	7

* Includes those with light perception.

** Education of those under age 33.

tomed to divide our cases into: 1. the totally blind from birth; 2. those blinded after 4–5 years, who apparently always have visual imagery through life and so have very different conceptions of the world; and 3. those with light perception, slight perception of objects, useful vision, etc." This quotation also justified the grouping by "Amount of Vision."

Merry (8, p. 156) states that "since the social situation is generally more normal, however, it may be inferred that those types of personality disturbances which are traceable directly to an institutional environment are not so frequent among pupils in these (day-school) classes." Thus a grouping of "Residential School Educated" vs. "Day-School Educated" was employed.

Finally, since employment status might have some repercussions in the personality of the blind, the groups were divided into "Employed" vs. "Unemployed" (excluding those in school).

The MMPI appeared suitable for this study because it had norms based upon a population comparable to that from which the sample of the blind was drawn, thus obviating the necessity for control groups. In addition, it has several scales based on psychiatric concepts and a large reservoir of items from which to construct other scales. Further, it has been found useful in the clinic (7) and for vocational guidance (4).

The American Printing House for the Blind printed the MMPI in grade one-and-a-half braille.[1] The directions were also transcribed into braille for each subject to read while they were read aloud. Noticing the difficulty of replacing the brailled cards into the box led to the modification of the instructions so that piles of "True," "False," and "Cannot Say" items were made. This facilitated the sorting, saving the subjects many minutes of useless fumbling.

In view of the empirical standardization of the MMPI as reported by Hathaway and McKinley (5), and contrary to the practice of most psychologists in giving tests to the blind, an unabridged inventory was used in an empirical attempt to determine whether or not the MMPI could be used with the blind. The items of the inventory, however, were

[1] Although grade two is the usual medium for books, grade one-and-a-half was chosen as the medium for the MMPI because grade one-and-a-half can be read by anyone who can read grade two, but the converse is not true. Usually the blind person first learns to read grade one, in which all the words are spelled out; then he learns grade one-and-a-half in which there are forty-four contractions; then grade two, having about a hundred and sixty-five contractions; the terminal grade, except for shorthand braille, is grade three which has still more contractions. For example, "confidence" is spelled out in grade one; in grade one-and-a-half, a symbol for "en" shortens the word by one sign; in grade two, signs for both "con" and "ence" shorten the word to five signs. Practically anyone who learns braille at all learns both grade one and grade one-and-a-half, but many people have neither the inclination nor the ability to remember the hundred and sixty-odd contractions of grade two.

submitted to a group of psychologists and others to secure their judgments as to the effect that lack of vision might have on the answers of the blind.

Item analyses were made, median profiles for each sex were constructed, significance ratios of the differences between means of the blind sexes and their appropriate norm groups and between the means of blind males vs. blind females were computed, and measures of the overlap[2] of the blind on the medians of the appropriate sex norm groups were determined.

Results and Discussion

Item analyses of the 550 items of MMPI reveal that the blind differ from the sighted norms significantly[3] on only 20 items. These are shown in Table 2.

On five items, the differences favor the sighted of both sexes, i. e., the sighted made a smaller number of deviant answers in the direction of significance than the blind. On two items the differences favor the blind males and the sighted females; on four items the differences favor sighted males and blind females; and on the remaining nine items the differences favor the blind of both sexes.

Item analyses of the 41 items chosen by a majority of the jury mentioned show that only four have significance ratios at the levels used when compared to their respective norm groups. These differences all indicate the blind to be better adjusted than the seeing. Thus, it is apparent that predictions of the jury were quite wide of the mark.

In considering the handicap of the blind, several directions in which the blind might be expected to deviate from the sighted suggest themselves. In the first place, the blind might be expected to be somewhat depressed since the existence of that condition limits their lives in ways that are early borne in upon them. This depression might be of the "smiling depression" variety, however. Secondly, one might expect the blind male to be more effeminate in his reactions since his handicap bars him from many of the pursuits open to the sighted man. The converse is not necessarily true because few of the activities considered differential in the direction of femininity are so affected. Thirdly, some tendency for the blind to be very careful about things would be expected, e.g., counting steps, crossing streets. In addition, Muhl (9) reports lack of confidence (even though obscured by a compensatory reaction of bravado), worry, and feelings of inferiority. Finally, Cutsforth (3) suggests a semantic

[2] Overlap is defined as number reaching or exceeding the median.

[3] Throughout the item analyses, the two per cent level of significance (significance ratio of 2.33) on either sex, and the five per cent level (significance ratio of 1.96) on the other sex, were used.

Table 2

Items on Which Both Sexes of the Blind Differ from the Sighted *

Significance Ratios**		Item	Scale
Males	Females		
+2.30	+2.80	My memory seems to be all right.	D, Sc, Pt
+2.30	+2.90	I very much like hunting.	Mf
+2.30	+4.10	I drink an unusually large amount of water every day.	Ma, Hy O-item
+2.40	+3.45	I am afraid when I look down from a high place.	Ma
+2.70	+2.50	The only interesting part of the newspapers is the "funnies."	F
+2.70	+2.10	Lightning is one of my fears.	No scale
+2.72	+2.98	Criticism or scolding hurts me terribly.	D
+3.40	+2.40	I would rather win than lose in a game.	L
+4.00	+2.62	At times my thoughts have raced ahead faster than I could speak them.	Ma O-item, Pd, Mf
+2.55	−2.40	My hardest battles are with myself.	Pd
+2.88	−3.50	I dread the thought of an earthquake.	No scale
−2.05	+4.00	Even when I am with others I feel lonely much of the time.	Pa, Pt, Sc
−2.80	+3.10	I like parties and socials.	No scale
−2.90	+2.90	I like to talk about sex.	Pd O-item, Mf
−2.92	+2.75	I am troubled by attacks of nausea and vomiting.	Hs, D, Hy, F
−2.05	−3.50	I have never been in love with anyone.	Sc
−2.15	−3.20	If I were in trouble with several friends who were equally to blame, I would rather take the whole blame than give them away.	No scale
−2.40	−2.80	I tend to be on my guard with people who are more friendly than I expected.	Pa
−2.45	−2.62	What others think of me does not bother me.	Hy, Pd
−2.65	−2.25	Children should be taught all the main facts about sex.	F

* At the two per cent level on one sex and at the five per cent level on the other.

** The +significance ratios indicate that the blind make fewer significant responses; the −significance ratios that the sighted make fewer significant responses.

difference probably exists between the blind from birth, or at a very early age, and the sighted. This factor might affect the reactions of the blind to the items making up an inventory of the type used here.

Construction of the median profiles of the total group of blind males and the total group of blind females reveals that apparent significant differences exist between the male blind of this group and the male norm group, but that no such differences exist between the female group and the

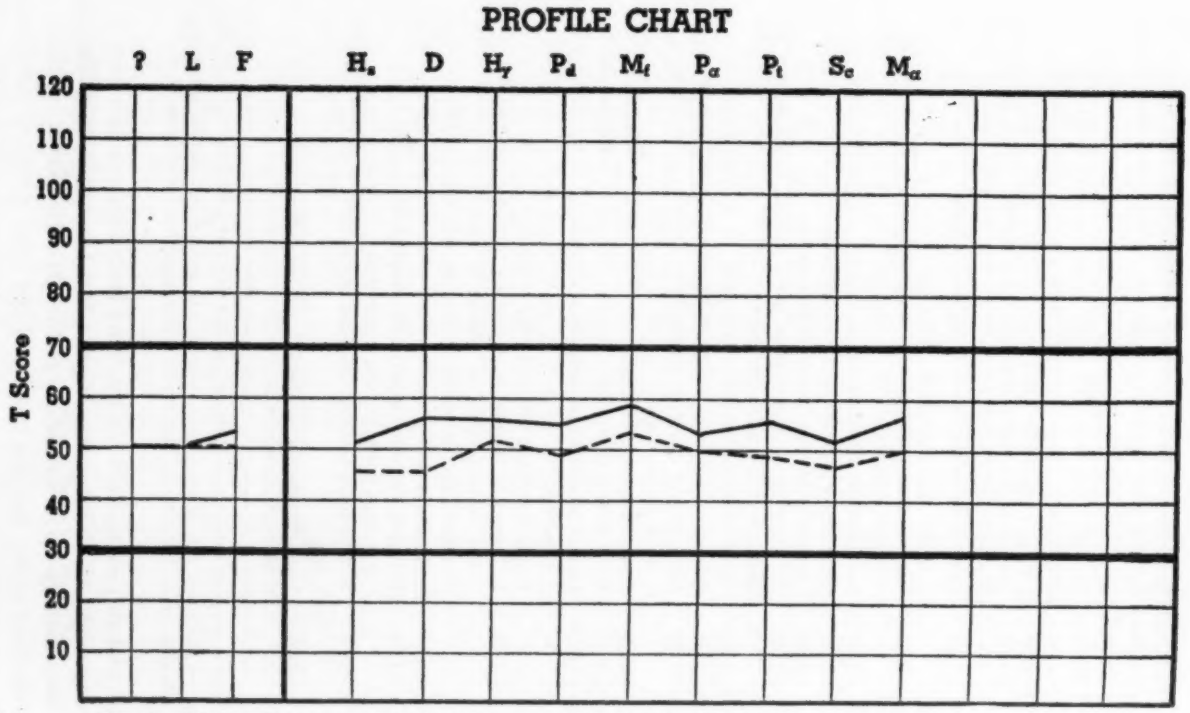

FIG. 1. Median T-score profile of total group of blind males (solid line) and median T-score profile of total group of blind females (broken line).

female norm group. The entire profile of the blind males is elevated, with only the validating scores, hypochondriasis, paranoia, and schizophrenia having median raw scores equivalent to T-scores of 53 or less. Fig. 1 illustrates these facts.

As a check of the validity of the profiles of the blind male and female groups, a compilation of the number of profiles showing one or more points above a T-score of 70 among groups of 100 males and 100 females picked from the norm groups and roughly equated to the blind samples as to age distribution shows that 23 per cent of the sighted males and an equal percentage of females thus chosen have such profiles. Of the blind samples under consideration, 24 per cent of the blind males and 16 per cent of the blind females have such profiles. These differences are not significant at the two per cent level. Since Mf scores do not appear on the profiles of the norm groups, this scale was not considered.

A further check on whether or not the profiles were valid results from a compilation of the percentage of T-scores above 70 which occur on each of the scales for both the norm groups and the blind groups. None of the differences found was significant at the two per cent level.

Thus, the evidence from the profiles appears to indicate that the test can be safely used with the blind without fear of exaggerating their personality traits. Two checks applied to the validity of the profiles confirm this conclusion, although the male blind group appears to be somewhat maladjusted. On the whole, the evidence clearly shows that the antecedent expectation that the blind, or an unusual proportion of the blind, would be found to be "abnormal" is not realized.

In examining the first of the points made above, the data gathered in this research show that, although the blind male is reliably more depressed than the sighted, the female is not. The blind male has a mean depression T-score of 58.1, a significance ratio between means of the male norm group and the sample of -3.20,[4] with an overlap of 88 per cent of the blind on the median of the norm group. The blind female, on the other hand, has a mean T-score of 49.6 on the depression scale, a significance ratio of $+0.060$ between the means, with an overlap of 48 per cent of the blind on the median of the norm group.

This difference may occur for one or more of the following reasons: first, the mores of our culture demand that the male earn a living for himself and, usually, a family. Since blindness distinctly handicaps the male in accomplishing this, depression is an expected result. A second explanation seems plausible. Hathaway and McKinley (6) found that a group of physically ill males confined to the hospital with no obvious psychiatric

[4] The $-$significance ratios indicate that the blind differ in the direction of abnormality; and the $+$significance ratios, that the sighted do.

FIG. 1. Median T-score profile of total group of blind males (solid line) and median T-score profile of total group of blind females (broken line).

complaint had a mean T-score of 60, while a similar group of females had a mean T-score of 55. It appears that this sample of blind males reacted to their blindness as though they were physically ill. However, the blind females did not react thus, suggesting that a third solution is possible, namely that the blind sample may not have been truly representative.

In regard to femininity of the males, the data reveal that the blind males are reliably more effeminate than the norm group. Their mean T-score on this scale is 60.6, the significance ratio between means of the blind sample and the male norm group is −5.99, and the overlap on the mean of the norms is 96 per cent (the mean was used because the median was not available). On this scale, the blind females did not differ significantly from the norm group of females.

On the scale which the authors of the inventory claim measures tendencies toward lack of confidence and compulsiveness, psychasthenia, the mean T-score of the blind male is 55.3, the significance ratio is −2.71, and 72 per cent of the blind sample reach or exceed the median of the norm group. As in both the scales above, the female blind do not deviate from their norm group.

As to the final hypothesis, the blind male differs from the sighted norms on the schizophrenia scale, the scale which comes closest to measuring the tendency to withdraw from the real world, to daydream, to have thoughts different from the normal. On this scale, the mean T-score of the blind males is 54.2, the overlap on the median is 76 per cent, and the significance ratio between the means is −1.95, not quite significant at the five per cent level. On the contrary, the female blind sample scores somewhat below the norms on this scale.

If we consider introversion to be most closely related to schizophrenia, the results recounted above contradict those of Brown (1), who found that the blind female is somewhat more introverted than the sighted female, and the blind male is somewhat more extroverted than the sighted male.

In addition to the scales mentioned above, the blind males prove to be maladjusted on the hypomania scale. Here the mean T-score was 56.1, the overlap on the median was 92 per cent, and the significance ratio between the means of the blind group and the male norm group was −3.38. If we consider extroversion to be most closely related to hypomania, this result confirms Brown's results with males (1).

Consideration of the significance ratios of the differences between the means of the total group of blind males and the total group of blind females reveals that the two groups differ at the two per cent level on four scales: depression (ratio of −2.98), masculinity-femininity of interests (ratio of −3.34), psychasthenia (ratio of −2.88), and schizophrenia (ratio

of −2.63). One other difference, that on the psychopathic deviate scale, appears nearly significant (ratio of −2.32). These differences all represent more normal reactions on the part of the female blind.

The hypotheses mentioned above suggest themselves as explanations of the differences found on the depression scale and might explain the differences found on the masculinity-femininity scale. However, the only explanation that holds for all the differences found between the blind and sighted males and between the two sexes of the blind is that of a sampling error. The fact that the author found it impossible to control the character of his sample of either sex, taking cases from the sources mentioned, supports this last hypothesis.

Computation of the significance ratios between the means of the pairs of groups in Table 1 fails to reveal any significant differences except between the means of blind males between the ages of 16–25 vs. those between the ages of 26–51 (significance ratio of 2.11 with the older group more maladjusted). Since this difference occurs in small groups (twelve vs. thirteen) from a sample of doubtful representativeness, it appears that this difference might disappear if the study had included a truly representative sample in each category. The evidence from the reactions of the female blind group tends to support this hypothesis, since no such differences occur.

Among the female groups compared, two significant differences occur. These differences confirm Merry's (8) statement to the effect that the more normal situation of the day-school class should result in better adjustment of the students in them. The day-school females prove to be better adjusted on both the schizophrenia scale (a significance ratio of 2.67, significant at the two per cent level), and on the hypomania scale (a significance ratio of 2.48, significant at the five per cent level).

Summary and Conclusions

The study was designed to determine whether or not the MMPI could be used in its present form with a population of blind people. The inventory, transcribed into braille, was administered to 50 blind people, 25 of each sex.

Item analyses and comparison of the profiles obtained indicate that valid results probably can be obtained. The fact that the median scores for the validating scales ("?," "L," and "F") do not differ to a noticeable degree from the norms also indicates that the results obtained are probably valid. Of the 550 items of the MMPI, item analyses reveal only twenty on which the blind differ significantly from the comparable norm group and on only five of these do both sexes of the blind answer in the maladjusted direction a significant number of times.

Comparison of the blind and the sighted on the various scales shows that the present sample of blind females is normal in its reaction to the MMPI. The males deviate from the norms to a marked degree (two per cent level) on the depression, masculinity-femininity of interests, psychasthenia, and hypomania scales.

A check of the percentage of abnormal profiles among the blind against those among the seeing reveals no significant difference. A second check, the percentage of abnormal scores on each scale for the norm group and the blind sample, also indicates that the blind are relatively normal.

Doubts as to the representativeness of he sample make it appear likely that differences might appear if it were better controlled, or that the differences obtained might disappear. The results of this research point to the fact that the MMPI in the braille edition probably can be used safely with the blind.

Received May 1, 1946.

References

1. Brown, P. A. Responses of blind and sighted adolescents to an introversion-extroversion questionnaire. *J. Psychol.*, 1938, **6**, 137–147.
2. Brown, P. A. Responses of blind and seeing adolescents to a neurotic inventory. *J. Psychol.*, 1939, **7**, 211–221.
3. Cutsforth, T. D. *The blind in school and society.* New York: Appleton, 1933.
4. Harmon, L. R., and Wiener, D. N. Use of the Minnesota Multiphasic Personality Inventory in vocational advisement. *J. appl. Psychol.*, 1945, **29**, 132–141.
5. Hathaway, S. R., and McKinley, J. C. A multiphasic personality schedule: I. Construction of the schedule. *J. Psychol.*, 1940, **10**, 249–254.
6. Hathaway, S. R., and McKinley, J. C. A multiphasic personality schedule: II. The measurement of symptomatic depression. *J. Psychol.*, 1942, **14**, 73–84.
7. Leverenz, Major C. W. Minnesota Multiphasic Personality Inventory: An evaluation of its usefulness in the psychiatric service of a station hospital. *War Med.*, 1943, **4**, 618–629.
8. Merry, R. V. *The education of visually handicapped children.* Cambridge: Harvard University Press, 1933.
9. Muhl, Anita M. Results of psychometric and personality studies of blind children at the California State School for the Blind. *Amer. Assoc. Instr. Blind, Proceed.*, 1930, 568–573.

Studies in Projective Techniques: I. The Reliability of a Multiple Choice Group Rorschach Test*

C. H. Lawshe, Jr. and Max H. Forster

Division of Applied Psychology, Purdue University

During the past few years Rorschach workers have conducted extensive research on the development of methods for group administration of the Rorschach test. Harrower-Erickson, a pioneer in this work, in an endeavor to obtain a minimum time-consuming group test which could be scored and evaluated without the aid of a trained Rorschach specialist developed the *Multiple Choice Test* (**4**) This test was designed primarily to serve as a screening test for filtering out the psychologically unfit, and it has been used as such by the military services, hospitals, schools, and industry with varying degrees of success. Articles in the current literature, such as those by Jensen and Rotter (**5**), Challman (**2**), Balinsky (1), Winfield (**9**), and Wittson, Hunt, and Older (**10**) strongly question the validity of this test as a screening instrument, but to date validation studies have not shown general agreement. This situation suggests the need for an investigation of the reliability of the test.

Purpose of the Study. The purpose of the study reported here was three-fold: 1. to test the reliability of an adaptation of the Harrower-Erickson *Multiple Choice Test;* 2. to analyze the test for characteristics which might permit scoring biases under certain conditions; and 3. to investigate the possibilities of increasing the reliability of the test through modified scoring based on item analysis.

Test Administration

A modified form of the Harrower-Erickson amplified *Multiple Choice Test* was administered to 655 college students, 371 men and 284 women, in four groups ranging from 100 to 200 students.

The testing room was an auditorium with 300 seats so arranged in tiers that vision from any part of the room was unobstructed. The auditorium was equipped with opaque shutters which allowed no light to enter the

*This article is a "prior publication," the authors paying complete costs. The scheduled 80 pages per issue is thereby increased by the corresponding amount, thus the "early publication" of this article is a direct contribution to the subscribers of the *Journal of Applied Psychology* without handicap to those authors whose articles are accepted and printed in their regular turn.

room. A standard lantern slide projector (Baloptican BDT, 500 watt lamp, 18 inch lens) was used to project standard Rorschach slides upon a flat white screen.

In place of the regular *Multiple Choice Test* answer sheet, a machine scored answer sheet (I.B.M. Form I.T.S. 10-00 A 189) was used. This answer sheet lists horizontally in each of ten rows digits from 1 to 10 in each of three vertical sections. These numbers represented for each inkblot exposure the ten standard answers which had been consecutively numbered for each group and placed on slides. First choice answers were indicated by marks above the numbers whereas alternate choices were marked below the numbers. The fact that multiple choice answers did not appear on the answer sheet but were projected upon the screen represents a departure from Harrower-Erickson standardization.

Each inkblot was projected on the screen three times for a period of thirty seconds each, during which time the auditorium was completely darkened except for the light from the slide projector. Following each inkblot exposure, the appropriate group of ten answers was shown on the screen for forty-five seconds during which time part of the auditorium lights were turned on while the subjects recorded their answers.

The standard introductory statements and directions were used except for such alterations needed to explain the use of the machine scored answer sheet.

Compilation of Basic Data

The 655 answer sheets were machine scored on all *poor* response items (Types 6-10). Answer types 8-10 were weighted double to obtain the .5 : 1 proportion for 6 and 7 vs. 8-10 type answers recommended by Harrower-Erickson (4). Separate scores for odd and even items of both first and alternate choices were recorded.

Each answer sheet was numbered and every fifth sheet was coded a "hold-out" to make up a control group for evaluation of scoring modification. Also on each subject's answer sheet was placed a sex code and an intelligence score as measured by the *Adaptability Test.*

All the basic data recorded on the answer sheets were punched on I.B.M. cards for machine tabulation and manipulation.

Reliability of Standard Scoring[1]

Since the primary purpose of this study was to determine the reliability of the Harrower-Erickson *Multiple Choice* (Group Rorschach) *Test,* the first step in the statistical analysis was to compute reliability coefficients for the standard scoring.

[1] Answer types 8–10 were given double weight.

Procedure. The split-half technique of measuring reliability was employed since the testing situation did not permit the use of the test-retest method. Three methods utilizing the split-half technique were considered: (1) the correlation of scores on odd numbered inkblots vs. those on even numbered inkblots; (2) the matching of scores for alternate inkblot exposures; and(3) the correlation of scores on odd numbered responses vs. those on even numbered responses.

A consideration of all the facts resulted in the selection of the odd-even response approach. To test the comparability of the two halves, so selected, means and standard deviations were computed. While the standard deviations (11.37 and 12.93) were not found to vary significantly, the means (11.7 and 18.3) were found to deviate to an extent not reasonably attributable to chance. Inasmuch as the split-half approach assumes the comparability of the two parts, and since this assumption is not completely fulfilled in the present instance, the results obtained must be evaluated in the light of this difference. While the correlations presented can probably be assumed to be spuriously low, the authors are of the opinion that the method employed has yielded the most nearly comparable halves obtainable from a random selection.

The odd item score vs. the even item score split-half technique was used for both first choices and alternate choices to compute reliability coefficients of correlation for the 655 cases. These correlation coefficients were stepped up by the Spearman-Brown prophecy formula to estimate the reliability of the full length test. Standard errors of the correlations and the standard error of a true score for first choices were calculated for the obtained reliability coefficients.

Results. The reliability coefficient for first choice scores was found to be $.61 \pm .01$ and for alternate choice scores $.40 \pm .01$. These results corroborate the findings of Wittson, Hunt, and Older (10) who report a reliability coefficient of .62P.E. $\pm$ .06 by the test-retest method.

As suggested by Guilford (3, p. 413) a reliability coefficient of .61 indicates that 61% of the scoring represents the true measure of the test, and 39% represents the proportion of the scoring attributable to errors of measurement. Since the coefficient of reliability is partially dependent upon the range of abilities of the group tested, the standard error of a true score was calculated. As an indicator of reliability for this population the $\sigma_{\infty 1}$ was found to be equal to 2.98 score points or 9.9% poor answers for first choice scores. Thus the estimated true score for an individual who obtains a score of 60% poor answers (critical score for screening) has 67 chances in 100 of falling between a 50.1% score and 69.9% score. Conversely there are 33 chances in 100 that the estimated true score (stated as a percentage score) will differ from the obtained score by more than 9.9 percentage points.

Sex Differences

In an analysis of the test for characteristics which might permit spurious results in scoring under certain conditions, sex differences and the intelligence factor were studied to determine relationships between these elements and test scores. As reported by Stavrianos (**8**, p. 168): "it has been well established by previous investigators that the Rorschach method is a useful instrument for the study of sex differences in children." This statement suggests the need for an analysis of the *Multiple Choice Test* for possible sex differences.

Procedure. Means and standard deviations were computed for male and female first choice scores. The standard errors of the differences for the two groups were calculated to determine the statistical significance of the observed differences.

Table 1

A Comparison of Means and Standard Deviations for Male and Female Scores (Per Cent of First Choices that are Poor Responses)

	Mean	Standard Deviation
Males	32.1	21.99
Females	26.9	17.59
Difference	5.2	4.40
Std. Error of Difference	1.547	1.094
Critical Ratio	3.36	4.02

Results. In Table 1 means and standard deviations for male and female first choice scores are compared. From these data it was determined from normal probability tables that the obtained differences in means and standard deviations of the two groups were significant beyond the 1% confidence level. Thus there is established evidence of true differences in male and female scores with the scores for men having a higher mean and greater variability. Item analysis, to be more fully discussed later, also reveals sex differences. This suggests the need for proper consideration of sex differences in score interpretations and points to the possible value of separate scoring keys for the two sexes.

Sex differences in scoring on the *Multiple Choice Test* might account for some of the disagreement found in validation studies in that some populations were predominantly men while others had a large proportion of women. Scoring allowances for sex differences might help eliminate adverse conclusions such as those reported by Winfield (**9**) from her study of screening a select group of women in military service.

Intelligence Factor

In consideration of the relationship between intelligence levels and scores obtained on the *Multiple Choice Test*, it was thought that possibly subjects of lower intelligence might not be as "test-wise" and would answer more naively without detection of poor answers as such.

Procedure. Intelligence scores as measured by the *Adaptability Test*, Form B, were correlated with scores made by the 655 subjects on the amplified *Multiple Choice Test.* These correlations were made with both first and alternate choice scores and with the number of alternate choice responses. The mean and standard deviation of the *Adaptability Test* scores were calculated to aid in the interpretation of the coefficients of correlation obtained.

Results. When the *Adaptability Test* scores were correlated with scores on the *Multiple Choice Test*, the following coefficients of correlation were obtained: with first choice scores $-.15 \pm .04$; with alternate choice scores $.001 \pm .041$; with number of alternate choices $.13 \pm .04$. The truth of the idea that persons of lower intelligence would tend to select more poor answers on this test was not proved. However, the strength of the intelligence factor in this test could be greater than these data reveal since the intelligence scores from a population of college students would tend to be restricted in variability and thereby yield a lower correlation coefficient. For these 655 students the mean score on the *Adaptability Test* was 24.3 with a standard deviation of 4.44 as compared to a mean score of 14.1 with a standard deviation of 7.26 for a general population.

Item Analysis

Preceding the establishment of scoring modifications, all 300 answers of the test were evaluated on their power to discriminate between subjects of good and poor criterion groups. Also item types predetermined by Harrower-Erickson (**4**) were studied for relative importance in the screening out of poor subjects.

Criterion Groups. The 131 control group cards which had previously been coded for separation were set aside, and the remaining 524 cards were then sorted on the basis of sex classification. The high and low 75 cards on first choice total score were selected from the male group and the procedure was repeated for the female group. From each of the high groups were selected the cards of the high 50 alternate choice scores, and correspondingly from each of the low groups the cards with the low 50 alternate choice scores were selected. The two male extreme groups were then combined with the female high and low groups to form the composite high 100 and low 100 criterion groups.

Item D-Values. An item count was made for the male, female, and

total groups to determine the per cent of the high and low criterion groups responding to each answer. After these percentages were established, the *D*-value nomograph (7) adaptation of the Kelley (6) technique for test item validation was employed to determine the discriminating power of each item. With the *D*-values thus established for each item, a tabulation by item types 1–10 was made for male, female, and total first choice responses to show the number of items of each type which fell at each *D*-value level.

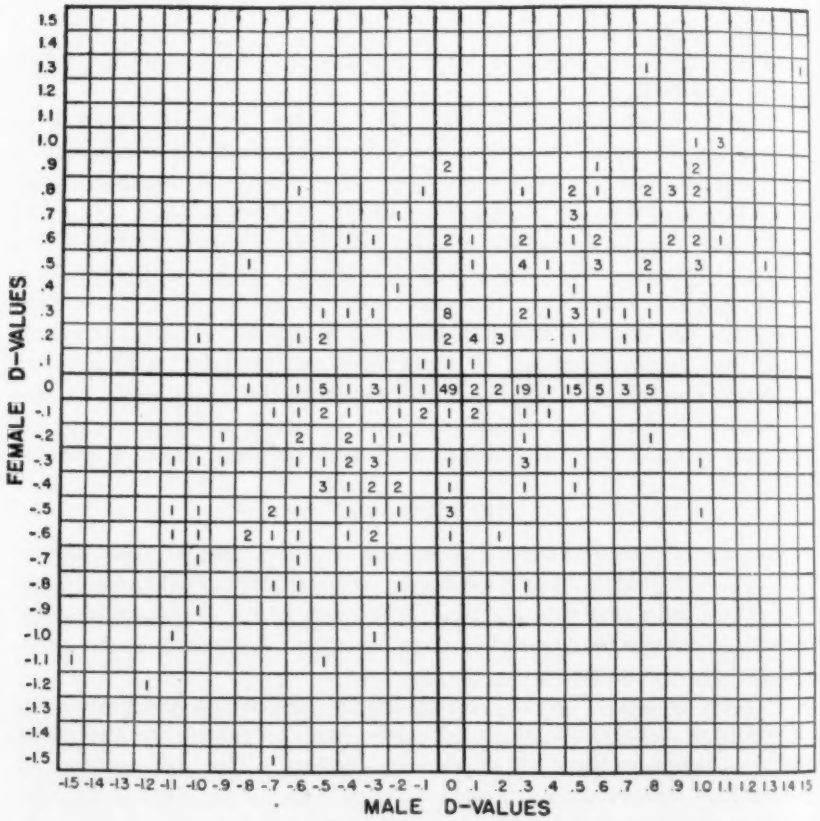

FIG. 1. Scattergram showing relationship between male and female *D*-values for the 300 items.

Male and female item *D*-values were plotted as a scattergram (Figure 1) on a quadrant to find the number of items tending to have the same discriminative power for both sexes and to note the general distribution of item *D*-values.

Results. Shown in Tables 2, 3 and 4, are the distribution of item *D*-values tabulated by answer type for first choice answers. As specified by Harrower-Erickson (**4**), answer types 1–5 are designated as *good* whereas types 6–10 are considered *poor* answers. From these tabulations can be seen the amount of overlap in the direction of the discriminative power of the *good* and *poor* answers. A plus item *D*-value indicates a higher percentage of answers from subjects of the high or *poor* criterion group,

Table 2

Distribution of *D*-Values by Good and Poor Item Types for 100 Men and 100 Women Composing Criterion Groups*

	Good Type Answers						Poor Type Answers					
D-Value	1	2	3	4	5	All	6	7	8	9	10	All
+1.4										1		1
1.3												
1.2												
1.1								2	2			4
1.0										1	1	2
.9								1	1	2	3	7
.8							2	2	1	5	4	14
.7								1	1	1	1	4
.6			2			2	4	2	2		1	9
.5			2			2	2	1	2	2	3	10
.4		1	1		1	3	2	1	2	8	1	14
.3		2	4		1	7	3	1	2	4	5	15
.2		1	7	4		12	3	6	3	6	4	22
.1		3	4	1		8	2	1	2			5
0	3		11	10		24	12	6	8	10	7	43
– .1	2	5	3	1		11						
.2	4	5	3	2	1	15						
.3	4	6	6	3	1	20						
.4	2	1	9	2		14						
.5	1	3	3	1		8						
.6	5	2				7						
.7	2	4	1	1	1	9						
.8	1	2				3						
.9	1	1				2						
1.0	1					1						
1.1						0						
1.2	1					0						
1.3			1			1						
Total	27	36	57	25	5	150	30	24	36	40	30	150

* Criterion groups: High group of 50 men and 50 women.
Low group of 50 men and 50 women.

Table 3

Distribution of *D*-Values by Good and Poor Item Types for 100 Men Composing Male Criterion Groups*

	Good Type Answers						Poor Type Answers					
D-Value	1	2	3	4	5	All	6	7	8	9	10	All
+1.5										1		1
1.4												
1.3										1		1
1.2												
1.1								2	2			4
1.0								1	1	6	4	12
.9							1			2	2	5
.8			3			3	1	2	1	3	3	10
.7			1		1	2	1		1	1		3
.6							3	1	1	2	5	12
.5		2	4	1	1	8	5	3	3	5	5	21
.4		1				1		1	1	1		3
.3	1		7	3		11	4	2	8	6	4	24
.2		1	4			5	1					1
.1		3	4	3		10		1				1
0	2	3	6	10		21	14	10	8	11	7	50
− .1	2	1	1			4				1		1
.2		4	4	1	1	10						
.3	2	5	5	3	1	16						
.4	4	2	2	2		10						
.5	1	4	8	1	1	15						
.6	3	3	4			10						
.7	3	1		1		5						
.8	3	1	1			5						
.9			2			2						
1.0	2	3	1			6						
1.1	3	1				4						
1.2	1					1						
1.3												
1.4												
1.5			1			1						
Total	27	35	58	25	5	150	30	24	26	40	30	150

* Criterion groups: High group of 50 men.
Low group of 50 men.

whereas a minus item *D*-value indicates a higher percentage of answers from subjects of the low or *good* criterion group.

Table 2 shows that for all subjects, 34 of the *good* answers (types 2–5) were selected by a greater percentage of the subjects of the high score or

poor group than of the subjects of the low score group. It should be noted that although the *poor* answers (types 6–10) in this table do not show any discriminative power in the opposite direction from what they were intended, the tables of male and female *D*-values (Tables 3 and 4) do give items of these types with negative *D*-values. This indicates a cancelling-out effect of negative *D*-values for *poor* answers when male and female an-

Table 4

Distribution of *D*-Values by Good and Poor Item Types for 100 Women Composing Female Criterion Groups*

	Good Type Answers						Poor Type Answers					
D-Value	1	2	3	4	5	All	6	7	8	9	10	All
+1.4												
1.3								1		1		2
1.2												
1.1												
1.0								1	2		1	4
.9							1		2		2	5
.8							2	2	2	3	4	13
.7		1				1	1	1	1			3
.6			3	1		4	2	2	2	3	2	11
.5	1	1				2	2	4	1	5	3	15
.4		1	2			3						
.3		3			1	4	3	4	2	5	1	15
.2		1	9	2	1	13	1	1				2
.1		1	1	1		3						
0	6	3	19	11	1	41	16	7	12	21	16	72
− .1	2	2	7	1		12			1			1
.2	1		6			7	1					1
.3	3	5	2	1	1	12	1	1	1	1	1	5
.4	3	4	3	2		12						
.5	4	3	1	2	1	11				1		1
.6	2	5	2	2		11						
.7	1	2				3						
.8		2		2		4						
.9	1					1						
1.0	1	1				2						
1.1			2			2						
1.2	1					1						
1.3												
1.4												
1.5	1					1						
Total	27	36	57	25	5	150	30	24	26	40	30	150

* Criterion groups: High group of 50 women.
Low group of 50 women.

swers are tabulated together for the calculation of *D*-values for all cases. As shown in the scattergram (Figure 1), there are wide discrepancies in item *D*-values for males and females; for example, answer No. 5 of Group C for Inkblot No. 6, "legs" (a type No. 4 answer) has a *D*-value of + .8 for females and − .6 for males. These *D*-values indicate that for females a greater percentage of the *poor* group selected the answer whereas for males the answer was selected by a greater percentage of the *good* group. The differences in male and female item *D*-values emphasize the existance of sex differences in scoring on this test as previously indicated by significant differences in means and standard deviations of male and female scores.

Further inspection of the item *D*-values for all subjects shows that about 22% of all the 300 answers have zero *D*-values. For male subjects alone, about 24% of the items have a *D*-value of zero; about 38% of all the items for females have *D*-values of zero. Table 5 lists those items having the most discriminative power in the intended direction.

Table 5

Items with Greatest Positive and Negative Discrimination Values

	Inkblot No.	Answer No.	Response	Answer Type	*D*-Value
Male	4	C-5	"Something squashed"	9	+1.5
	9	C-5	"Men's faces on the sides"	3	−1.5
Female	4	C-5	"Something squashed"	9	+1.3
	7	A-1	"Smoke"	7	+1.3
	4	B-6	"A man sitting down"	1	−1.5
Total	4	C-5	"Something squashed"	9	+1.4
	9	C-5	"Men's faces on the sides"	3	−1.3

It is of interest to note from the tabulations of *D*-values by item types (Tables 2, 3, 4) that *quantitatively speaking* type No. 10 does not appear to be any more significant than the other *poor* types in its power to separate the *poor* from the *good* subjects.

Modified Scoring

Modifications in scoring were developed in an endeavor to raise the reliability of the test. These modifications were aimed at the scoring of only the more discriminative items and the development of separate scoring keys for males and females.

Procedure. Seven scoring modifications were developed and, from the scores of the control group of 131 cases, reliability coefficients were calculated for each system by the odd even split-half technique. By inspection of the tabulations of item *D*-values (Tables 2, 3, 4) *D*-value cut-

off points for *poor* item types 6–10 were determined for the first five modified scorings. In the sixth and seventh scoring modifications both *good* answers (types 1–5) with negative *D*-values and *poor* answers (types 6–10) with parallel positive *D*-values were scored; each *good* answer was given a minus value and each *poor* answer was scored as a plus value. To get a total score for each individual the algebraic sum of the plus and minus items was obtained and a constant value of twenty was then added to make all scores positive for ease in machine tabulation in calculation reliability coefficients.

Modification No. 1. *D*-values of + .4 or higher (65 items)
Modification No. 2. Female *D*-values of + .4 or higher (Female key of 53 items)
Modification No. 3. Male *D*-values of + .4 or higher (Male key of 72 items)
Modification No. 4. Female *D*-values of + .3 or higher (Female key of 68 items)
Modification No. 5. Male *D*-values of + .3 or higher (Male key of 96 items)
Modification No. 6. Female *D*-values of + .3 or higher for 6–10 type answers and − .3 or lower for 1–5 type answers (Female key of 128 items)
Modification No. 7. Male *D*-values of + .3 or higher for 6–10 type answers and − .3 or lower for 1–5 type answers (Male key of 171 items)

Results. As shown by the reliability coefficients in Table 6 when the less discriminative items were removed from the scoring, there were not enough items left to raise or even maintain the reliability coefficients

Table 6

Stepped Up Reliability Coefficients
(First Choice Scores)

	Standard Scoring	Modified Scoring						
		1	2	3	4	5	6	7
Male	.583	.426	xxx	.532	xxx	.586	xxx	.694
Female	.648	.541	.544	xxx	.536	xxx	.594	xxx
Both	.606	.478	xxx	xxx	xxx	xxx	xxx	xxx
No. of Items	150	65	53	72	68	96	128	171
D-Value Cut-off	*	+.4	+.4	+.4	+.3	+.3	±.3	±.3

* Range: Male −.6 to +1.5
Female −.5 to +1.5
Both 0 to +1.5

originally obtained except for two male keys, Modifications No. 5 and No. 7. It should be noted that in Modification No. 5 with the lowering of the critical *D*-value from + .4 (Modification No. 3) to + .3, twenty-four items were added to the scoring. Also in Modifications No. 6 and No. 7 the number of discriminative items scored was approximately double that of the other modified scorings at the same *D*-value level and marked rises in reliability coefficients resulted. The apparent reason Modifications No. 5 and No. 7 gain the higher levels of reliability lies in the fact that fewer items are eliminated than are in the total and female modifications with the same *D*-value cut-offs. The male scoring of *poor* answers alone at the + .3 *D*-value cut-off level eliminates 36% of the originally scored items whereas for the total and female scorings 46.7% and 54.7% respectively were eliminated. These findings strongly indicate the need for more discriminative items in the test in order to improve its reliability. If, however, item validity is not increased, more desirable test reliability might be gained through the inclusion of more inkblots with multiple choice answers of discrimination power at least equal to the answers now in use.

Summary and Conclusions

Reliability coefficients were computed for the Harrower-Erickson *Multiple Choice* (Group Rorschach) *Test* by the split-half technique from the scores of 655 college students. By this method the coefficients were found to be .61 ± .01 for first choice scores and .40 ± .01 for alternate choice scores. Sex differences in responses, the strength of the intelligence factor, and the item discrimination values were studied for indications of needs for scoring modifications. Seven scoring revisions were developed and tested for reliability on a control group of 131 subjects.

The results of this study support the following conclusions:

1. The low reliability coefficients obtained from scores of college students on the *Multiple Choice Test* strongly indicate the unreliability of the standard scoring and reflect the weakness of the test as a screening instrument. However, in cases where a low selection ratio can be used, that is, when the object is to select few individuals from a large number of candidates, assuming validity comparable to the obtained reliability the test might be effectively used for selection.

2. Significant sex differences in score distribution and item selection reveal the need for separate male and female scoring keys.

3. For the restricted population studied, the intelligence factor does not seem to contribute significantly to the variabilities in scores on the *Multiple Choice Test*.

4. The reversal of the intended discriminative power of some answers indicates a need for a review of the original classification of items.

5. The insufficient number of strongly discriminative items in this test does not permit a significant raising of the reliability through the elimination of weak items. The suggestion is advanced that an increase in the number of inkblots might render the test more useful in individual applications.

Received January 27, 1947.

References

1. Balinsky, B. The Multiple Choice Group Rorschach Test as a means of screening applicants for jobs. *J. Psychol.*, 1945, **19**, 203–208.
2. Challman, R. G. The validity of the Harrower-Erickson Multiple Choice Test as a screening device. *J. Psychol.*, 1945, **20**, 41–48.
3. Guilford, J. P. *Psychometric methods.* New York and London: McGraw-Hill Book Company, Inc., 1936.
4. Harrower-Erickson, M. R., and Steiner, M. E. *Large scale Rorschach techniques.* Springfield, Ill.: Charles C. Thomas, 1945.
5. Jensen, M. B., and Rotter, J. B. The validity of the Multiple Choice Rorschach Test in officer candidate selection. *Psychol. Bull.*, 1945, **42**, 182–185.
6. Kelley, T. L. Selection of upper and lower groups for the validation of test items. *J. educ. Psychol.*, 1939, **30**, 17–24.
7. Lawshe, C. H., Jr. A nomograph for estimating the validity of test items. *J. appl. Psychol.*, 1942, **26**, 846–849.
8. Stavrianos, B. An investigation of sex differences in children as revealed by the Rorschach method. *Rorschach Res. Exch.*, 1942, **6**, 168–175.
9. Winfield, M. C. The use of the Harrower-Erickson Multiple Choice Rorschach Test with a selected group of women in military service. *J. appl. Psychol.*, 1946, **30**, 481–487.
10. Wittson, C. L., Hunt, W. A., and Older, T. J. The use of the Multiple Choice Group Rorschach Test in military screening. *J. Psychol.*, 1944, **17**, 91–95.

A Note on the Hankes Test Scoring Machine*

E. K. Strong, Jr.

Stanford University

and

E. J. Hankes

Engineers Northwest, Minneapolis

This is a device to score tests such as the Strong Vocational Interest Test. All that the operator has to do is to insert a special answer sheet in the machine and turn a switch successively to 42 positions and read the total scores on the 42 scales. Speed of scoring is the speed with which 42 scores can be read and recorded by the operator plus the time of inserting the answer sheet in the machine. Actually 40 men's blanks can be scored an hour.

The special answer sheet used in the machine is a piece of foil backed paper, $8\frac{1}{2} \times 11$ inches in size, in which holes have been punched corresponding to the 1,200 possible responses to the 400 items on the Vocational Interest Blank. A clerk can punch 12 of these answer sheets an hour.

Present unit costs for scoring the men's blank on 42 scales depends upon the way in which the man answered the test. If he marked a test booklet the unit cost is about 2 cents per scale; if he marked an IBM answer sheet the unit cost is slightly less; if he marked an answer sheet designed by Hankes and distributed by the Stanford University Press, the unit cost is still less. The difference in costs reflects the ease of punching holes in the special answer sheet used in the machine from the responses on the test itself or from the answer sheet used by the person taking the test.

It will be recalled that the Strong test has 400 items to which the person taking the test responds in one of three ways (Like, Indifferent and and Dislike) giving a total of 1,200 possible responses. Each of these responses has a fixed value in each of 42 occupations or Keys. This gives a total of $42 \times 1{,}200$ or 50,400 weights ranging from $+4$ to -4. For

* This article is a "prior publication," the author paying complete costs. The schedule 80 pages per issue is thereby increased by the corresponding amount, thus the "early publication" of this article is a direct contribution to the subscribers of the *Journal of Applied Psychology* without handicap to those authors whose articles are accepted in their regular turn.

example, liking item No. 2 has a value of + 4 for life insurance salesman, + 2 for mortician, − 2 for engineer, etc.

The fixed grill scoring machine combines all the varying weights (+ 4 to − 4) for the 400 responses among a possible 1,200 responses on each of the 42 Keys and as the control switch is turned successively to the 42 Keys, the total score is given for each of the Keys. How is this accomplished?

In the fixed grid system, electrical resistance units are used to correspond to the weights (+ 4 to − 4) being counted, i.e., for a weight of 1 a resistance of value "R" capable of passing one unit of current is used; for a weight of 2 a resistance of value ½R is used, passing 2 units of current, etc. These resistances are arranged on cards, the cards corresponding to the 1,200 possible answers in the test. On each card there are 42 resistance units corresponding to the 42 Keys. These are distributed on the card so that those giving plus weights are above the center of the card, those giving minus weights below the center of the card. Along the center of the card there is a wire connecting all the resistances together. This wire is also connected to a contact pin on the board which comes in contact with the special answer sheet. If there is a hole in the special answer sheet the pin is activated, otherwise not.

The current flowing in the activated resistors is collected on each of 42 cross wires, each one corresponding to a particular occupation. These 42 wires are connected to a 42-positioned switch by means of which the operator can select any particular occupation for measurement. When the switch is set to any given occupation, all the currents flowing in the resistors representing the score values of that occupation are fed into an evaluator is an electrical circuit similar to those widely used during the war on flight control, fire control, and other precise control or measurement problems requiring instantaneous solutions.

The evaluator is a bridge circuit, one side of which consists of a pair of precisely calibrated resistors through which the current from the "occupation" resistors flows. This double resistor is so arranged that it acts as an adding machine, adding the plus and minus counts together to give the total raw score. The remainder of the bridge circuit is arranged with an automatically driven motor controlled by an electronic amplifier. The evaluator is known technically as a "self-balancing bridge circuit." The motor drives a potentiometer which is connected into the bridge circuit. This motor driven potentiometer brings the bridge into balance when the current flowing through the potentiometer corresponds to that flowing through the accurately calibrated resistors. In this way a dial connected to the same motor shaft as the potentiometer can be calibrated to read the raw score.

The use of the 42-positioned switch permits a rapid reading of the scores with only one insertion of the test blank.

The machine has been in operation since February 1946 and has amply demonstrated its value both in markedly decreasing the labor and greatly reducing the cost of scoring the Vocational Interest Blank. Comparison of scores by use of the machine and of the IBM Tabulator discloses a variation between the two of only 1 to 2 standard scores.

Received January 31, 1947.

Book Reviews

Oberdahn, R. C. *How to select foremen and supervisors.* Deep River, Conn.: National Foremen's Institute, Inc., 1944. Pp. 53. $2.00.

The author of this book has engaged in slection and training work for many years and he presents his experiences to help other industrialists formulate a selection program for foremen.

His procedure, very briefly, is to analyze the foremen's jobs, select tests for experimental purposes, establish minimum scores on those tests that correlate with superintendent's ratings of the existing foreman force, and incorporate those tests into the reviewing and character investigating procedures used in selecting foremen. This method is to be substituted for the widely used practise of appointing the best operator to a foremanship vacancy.

Oberdahn writes that he does not wish "to be set up as a dogmatic expert in the field of psychology" and that "some criticism as to the accuracy of this method may come from the professional psychologist." It is quite unlikely, however, that industrial psychologists will criticize his general approach.

In technical details he is not above reproach. For example, he uses the terms "degree of intelligence," "high degree of intelligence" and "I Q" interchangeably. He claims to have found his results to be both valid and reliable although these terms are usually applied to methods rather than to results. He presents no data pertaining to reliability and validity. His procedure of testing the tests against the existing foreman force is an accepted practice in installing test batteries but suffers, in his example and probably also in most industrial concerns, from an extremely small sample when working at the foreman level.

Oberdahn analyzes foremen's jobs into three broad categories; technical aspects, human relations aspects, and those aspects involving a combination of these two. He breaks each of these broad groups in more specific job duties and indicates which of three psychological factors are involved in each. These three factors are intelligence, temperament, and aptitude (meaning mechanical aptitude). He does not present any data to justify these relationships of job duties to psychological factors, but he does correctly state that the important thing is to experiment with the tests rather than to accept their names at face valus.

He suggests giving, on an experimental basis, all of the tests of the three psychological factors that are available to the experimenter. He

has used, apparently on one foreman force, four intelligence, four reading, four arithmetic, seven "temperament," and five mechanical aptitude tests. It is doubtful to the reviewer if so many tests need be used even for purposes of testing them. It is also doubtful if all of these tests would be found useful outside of a few fabricating industries.

This book was written for industrialists and not for psychologists. The latter will find very little of value in it. It may prove valuable in indicating to industrial executives the need for and value of modern selection techniques.

Harold F. Rothe

Stevenson, Jordan, and Harrison, Inc.,
Chicago, Ill.

Benson, Barbara Elna, *Music and sound systems in industry.* New York: McGraw-Hill Book Company, Inc., 1945. Pp. ix + 124. $1.50.

For the last odd hundred years the production engineer has been waging an unceasing struggle against the "physical" fatigue of industrial workers. In terms of caloric expenditure, only a very small percentage of jobs in modern industry can be classified as "heavy work." However, the increasing mechanization and specialization of industrial operations brought to the production man a problem-child of its own—that of boredom-fatigue

This type of fatigue has an important psychological component and it is therefore logical to combat boredom-fatigue by psychological means. The use of "music-while-you-work" appears to be a promising approach although it is not—by any stretch of imagination or enthusiasm— a panacea for all the "labor troubles." The interest in music for industry, and in industrial broadcasting in general, has been growing rapidly in the last five to ten years and was noticeably stimulated by the war-time search for means with which to improve the work adjustment of the individual and to increase production. There was a job to be done but no trained personnel for it. The book reviewed has been written as a *vade-mecum* for persons assuming responsibility for creating and directing an industrial broadcasting program. The manual is avowedly practical and treats the subjects related to the organization and operation of an industrial broadcasting system on an elementary level.

The activities of a broadcast operator include not only selection and broadcasting of musical programs but also preparation and presentation of newscast, special announcements and, frequently, the service job of paging. The music broadcast in industry falls into four categories: 1) pre- and post-work music, 2) work music, 3) lunch music, and 4) special-events music, such as designed for holiday programs and personal announcements on anniversaries and birthdays. Programming of work

music is a difficult and responsible task although *all* the aspects of industrial broadcasting are relevant to the building and maintenance of morale of the industrial worker.

Music played during work has for its primary goal relief from the monotony of work. The author is opposed, and rightly so, to the idea that one can aim at increasing the production rate by playing music of a fast tempo and a marked rhythm to which the workers would "keep time." Music played during work should be esthetically satisfying yet not distracting. An industrial plant is not a music hall and industrial music has its own criteria of "goodness," a definite melody line which can be easily followed and remembered, a steady rhythm, and a relatively constant volume. The last principle is particularly important when music is broadcast in areas of a high noise-level. Many of the commercial recordings are unsatisfactory from one or another points of view just mentioned. A novice may find, therefore, quite useful the forty pages devoted to discography, a classification of commercial records according to their potential industrial use.

The broadcasting schedule should be based on the study of the peculiarities of the individual plants, especially the daily output curves. In the absence of such data, music may be played for thirty minutes at about the middle of each half of the shift, a time when the production curves are apt to exhibit a "boredom slump." The reviewer shares the author's concern about exposing the worker to too much music which would defeat completely the purpose for which industrial music has been introduced.

The section on scientific research on music in industry is one of the less satisfactory chapters but it does show the layman some of the techniques and criteria for the evaluation of industrial broadcasting, such as production records and music-attitude inventories. The author makes a plea for intensified scientific investigation of the problem and raises a number of questions confronting the director of an industrial broadcasting program, some of which represent fruitful research topics.

Josef Brozek

Laboratory of Physiological Hygiene,
University of Minnesota

Garrett, Annette. *Counseling methods for personnel workers.* New York: Family Welfare Association of America, 1945. Pp.187. $2.00.

The primary purpose of this book is a statement of basic counseling principles which will be of practical value to industrial counselors.

The first few chapters deal with fundamental principles underlying human behavior, the material being presented in non-technical terminology and in such a clear and simple manner that it can be understood by

persons untrained in the field of psychology. This has been accomplished without falling into the errors of oversimplification. Rationalization, repression, displacement, projection, ambivalence, and other mechanisms are discussed in such a way as to be fascinating reading to any person who is average or above in mental ability. The author succeeds well in establishing the viewpoint that such mechanisms are commonly found in every-day life and are normal rather than abnormal. This material is presented in an exceedingly practical manner and should be required reading for all industrial executives as well as counselors.

Basic counseling methods (including interviewing, understanding the problem and handling the problem) are covered in the second part. The chapter on interviewing is inadequate if read by itself, but is supplemented by supporting material appearing throughout the book. A third section covers specific counseling problems such as induction and exit interviews, in-plant and out-plant adjustment, referrals, and working with foremen.

The final chapters cover the relationships between counseling and case work; and are in the nature of propaganda intended to secure greater recognition, appreciation, and use of community agencies by industrial concerns. Case workers will highly applaud this section, but the typical industrial counselor and supervisor will react quite differently because, unfortunately, most such persons wince when any relationship between their work and case work is mentioned. The reviewer believes that omission of this material would have resulted in counselors generalizing on the practicality of the book to include practicality of caseworkers and agencies, thus resulting in greater cooperation than will be obtained through the more direct sales methods presented in these chapters.

With regard to the directive versus non-directive controversy in counseling, this book takes a middle of the road position. The author stresses the dangers and inadvisability of giving advice, but recommends giving information and warns against too much listening and too little action. Directive counselors are apt to consider the book too non-directive, and non-directive counselors will condemn it for being too directive. Counselors belonging to neither of the extreme schools will praise the book for having a refreshing eclectic viewpoint.

The book is filled with numerous cases taken from industrial records. These are well selected to emphasize and clarify the author's arguments, and will do much to make the book usable to industrial counselors. The author's vocabulary is reasonably adequate when quoting industrial cases, in spite of a few instances where the wording is radically different from what would be found in industry; for example, referring to a disciplinary lay-off without pay as "he had been requested to take a week off." In

view of the author's background, such lapses are not surprising; it is surprising only that there are not a great many more.

This book will be valuable to all industrial and union personnel who are directly or indirectly concerned with counseling. It will also be valuable to experienced case workers who lack understanding of industrial counseling and to new workers in the case field. Personnel courses offered to persons having little or no formal training in psychology can well use the book for supplementary reading if it is desired that such students gain a maximum amount of general information about industrial counseling with a minimum expenditure of time and effort. The book was not intended for a psychology text or reference, and would have little value for such use; although many industrial and academic psychologists could profit from the manner in which the book treats a professional field in such way as to be fascinating, understandable and valuable to people in business and industry.

Clifford E. Jurgensen

Minneapolis Gas Light Co.

Wilson, Frances M. *Procedures in evaluating a guidance program.* New York: Bureau of Publications, Teachers College, Columbia University, 1945. Pp.210. $2.60.

The purpose of this book, according to the author, is five fold:

1. To state the functions of guidance, to offer criteria for determining the success or failure of a guidance program in terms of those functions, and to consider the factors which contribute significantly to that success or failure.
2. To summarize those evaluation studies which have already been made in the field of guidance.
3. To describe a program of evaluation for use in surveying the guidance program of a secondary school system in a large city.
4. To criticize the techniques of observation, interview, questionnaire, and rating scale, with special reference to their use in evaluating guidance programs.
5. To develop a program of self-evaluation for use by an individual school.

The statements regarding the functions and philosophy of guidance are good. In the discussion of criteria, some 154 questions are presented, one following the other, which the author proposes to be answered in the evaluation of a secondary school program. Just how the answers may be put together to form an over-all estimate is not clear, or how much significance should be assigned to any particular item or group of items is not indicated.

The listing of factors which must be considered in respect to a guidance program should be studied carefully by those who intend to introduce a program. Even those persons who are at present administering a successful program might find it to their advantage to go over this list and relate it to their own program. The summary of evaluation studies tends to concentrate on the New York School Survey whereas other studies are briefly mentioned or omitted entirely. In general the book attemps to cover too much ground in 210 pages.

L. E. Drake

University of Wisconsin

Link, Henry C., and Harry Arthur Hopf. *People and books: a study of reading and book-buying habits.* New York: Book Industry Committee, Book Manufacturers' Institute, 1946. Pp. 167. $10.00.

Would the war boom in book buying continue, decline gradually, or collapse during the reconversion period? The book industry needed an answer to this question in order to protect its heavy investment in plants and organization. With the financial backing of the entire industry, the Psychological Corporation and the Hopf Institute of Management were commissioned to conduct a consumers' survey upon which estimates of future market trends could be based. *People and books* is a report of the findings of this study, made in the spring of 1945, by the men who served as joint directors.

Conventional public opinion polling techniques were used in gathering data. These are fully described. The population interviewed consisted of 4,000 individuals, fifteen years of age and older, representing the non-farm, civilian population of the United States. Interviews were also held with 225 book dealers and distributors and with 100 college and university administrators.

The main body of the report consists of a series of simple tables and graphs, showing percentages of people in the consumers group who answered the questions asked according to each of various alternatives, with accompanying text describing and interpreting these statistics. Basic breakdowns are made according to recency of reading, income level, education, age, sex, and religious background. The questions deal with such matters as frequency of reading, types of books and subjects read and preferred, physical characteristics and price of last book read, where and how books are obtained, book ownership, price preferences, how and why books are selected, time spent in reading as compared with other activities, comparison of recent with estimated future book reading and buying. "Correlations" are reported between some of these variables, apparently from inspection of the percentage data, but no coefficients of correlation are given.

The major conclusion is that "everything in our survey points to a long-term gain in the reading, and therefore in the purchasing, of books." The validity of this inference might be questioned. Years of formal education appear to be more closely related to frequency of readership than any other variable, although there is some relation between frequency of reading and socio-economic status. The authors predict a pronounced trend toward increased reading of non-fiction. They also anticipate the creation of a huge market for low-priced books, side by side but not in competition with the existing market for higher-priced books. They conclude that people will buy books that interest them, without too much regard for price. Does this account for the $10 price set on their own volume!

Although some adjustment was made in the sampling quotas, no data were gathered on the reading of the millions of American men who were in military service at the time the survey was made. Nor is there any referenec in the interpretation of the findings to relevant information that has been gathered by other investigators about this large and influential group of readers. No effort has been made to relate the findings of this survey to some of the other studies of "people and books" that have been made by highly competent investigators. The book is attractively printed and bound, but it contains neither bibliography nor index.

Alice I. Bryan

Columbia University

New Books, Monographs, and Pamphlets

Books, monographs, and pamphlets for listing and possible review should be sent to Donald G. Paterson, Editor, Department of Psychology, University of Minnesota, Minneapolis 14, Minnesota

The mind and death of a genious. David Abrahamsen. New York: Columbia University Press, 1946. Pp. 228. $3.00.

Non-projective personality tests. Harold A. Abramson, et al. New York: New York Academy of Sciences, 1946. Pp. 147, $1.75.

Tutoring as therapy. Grace Arthur. New York: The Commonwealth Fund, 1946. Pp. 125. $1.50.

Industrial peace: a progress report. Wilson Baden. New York: The National Association of Manufacturers, 1946. Pp. 104.

An approach to guidance. Edna D. Baxter. New York: D. Appleton-Century Co., 1946. Pp. 305. $2.50.

A psychology of growth. Bert I. Beverly. New York: McGraw-Hill Book Co., Inc., 1946. Pp. 235. $2.50.

Psychology, normal and abnormal. James W. Bridges. New York: Pitman Publishing Corporation, 1946. Pp. 470. $4.50.

Dynamics of learning. Nathaniel Cantor. Buffalo: Foster & Stewart, 1946. Pp. 282. $3.00.

The American Transit motor ability test. Glen U. Cleeton, et al. New York: American Transit Association, 1946. Pp. 12. Gratis.

The personal reaction test for transit employees. Glen U. Cleeton, et al. New York: American Transit Association, 1946. Pp. 7. Gratis.

The standard examination for transit employees. Glen U. Cleeton, et al. New York: American Transit Association, 1946. Pp. 8. Gratis.

Item-analysis data: their computation, interpretation, and use in test construction. F. B. Davis. Cambridge: Harvard Graduate School of Education, 1946. Pp. 42. $.75.

Learning to use hearing aids. Arthur I. Gates and Rose E. Kushner. New York: Bureau of Publications, Teachers College, Columbia University, 1946. Pp. 77. Limited free distribution.

Human genetics. Two volumes. R. Ruggles Gates. New York: The Macmillan Co., 1946. Pp. 1518. $15.00.

Encyclopedia of psychology. Philip L. Harriman, et al. New York: The Philosophical Library, Inc., 1946. Pp. 897. $10.00.

The measurement of understanding. Nelson B. Henry, Editor. Chicago: The University of Chicago Press, 1946. Pp. 338. $2.25.

Vocational personalities. Goldie R. Kaback. New York: Bureau of Publications, Teachers College, Columbia University, 1946. Pp. 116. $2.10.

Selection and employment of transit operating personnel. Merwyn A. Kraft, Robert F. Royster, and Glen U. Cleeton. New York: American Transit Association, 1946. Pp. 51. $3.50.

Aging successfully. George Lawton. New York: Columbia University Press, 1946. Pp. 266. $2.75.

Handbook of correctional psychology. Edited by Robert M. Lindner and Robert V. Seliger. New York: Philosophical Library, Inc., 1946. $10.00.

Manual of diagnostic psychological testing. David Rapaport, Roy Schafer, and Merton Gill. New York: The Josiah Macy, Jr. Foundation, 1946. Pp. 104. $.75.

The psychological problems of religion. I. Ritual: psychoanalytic studies. Theodor Reik. New York: Farrar, Straus and Co., Inc., 1946. Pp. 367. $5.00.

Modern clinical psychology. T. W. Richards. New York: McGraw-Hill Book Co., Inc., 1946. Pp. 331. $3.50.

Effective study. Francis P. Robinson. New York: Harper & Brothers, 1946. Pp. 262. $3.00.

White caps. The story of nursing. Victor Robinson. Philadelphia: J. B. Lippincott Co., 1946. Pp. 425. $3.75.

The placement interview for transit employees. Robert F. Royster, Glen U. Cleeton, and Merwyn A. Kraft. New York: American Transit Association, 1946. Pp. 45. $3.00.

High-school credit and diplomas through examinations and out-of-school experiences. David Segel. Washington, D. C.: Superintendent of Documents, Government Printing Office, 1946. Pp. 46. $.20.

A simplified guide to statistics. G. Milton Smith. New York: Rinehart & Co., Inc., 1946. Pp. 109. $1.25.

Their mothers' sons. Edward A. Strecker. Philadelphia: J. B. Lippincott Co., 1946. Pp. 220. $2.75.

Selective job placement. Tobias Wagner. New York: The National Conservation Bureau, 1946. Pp. 151. $2.50.

Psychiatric interviews with children. Helen L. Witmer. New York: The Commonwealth Fund, 1946. Pp. 443. $4.50.

Personnel classification tests. War Department Technical Manual. Washington, D. C.: Superintendent of Documents, Government Printing Office, 1946. Pp. 90. $.30.

Yivo annual of Jewish social science. New York: Yiddish Scientific Institute—Yivo, 1946. Pp. 319.

Safety education in the schools. New York: Board of Education, 1946. Pp. 48. $.20.

Where to find vocational training in New York City: a directory. New York: Vocational Advisory Service, 1946. Pp. 107. $2.00.

tific Ins-

ion, 1946.

y. New